



日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 3 年 1 0 月 6 日
Date of Application:

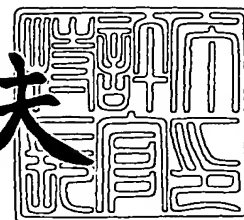
出 願 番 号 特 願 2 0 0 3 - 3 4 6 6 0 8
Application Number:
[ST. 10/C]: [J P 2 0 0 3 - 3 4 6 6 0 8]

出 願 人 株式会社日立製作所
Applicant(s):

2 0 0 3 年 1 0 月 2 3 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫





【書類名】 特許願
【整理番号】 NT03P0748
【提出日】 平成15年10月 6日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 12/00
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 芹沢 一
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 下藺 紀夫
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 山本 康友
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 岩見 直子
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社日立製作所
【代理人】
 【識別番号】 100068504
 【弁理士】
 【氏名又は名称】 小川 勝男
 【電話番号】 03-3661-0071
【選任した代理人】
 【識別番号】 100086656
 【弁理士】
 【氏名又は名称】 田中 恭助
 【電話番号】 03-3661-0071
【選任した代理人】
 【識別番号】 100094352
 【弁理士】
 【氏名又は名称】 佐々木 孝
 【電話番号】 03-3661-0071
【手数料の表示】
 【予納台帳番号】 081423
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

ストレージ装置が持つある記憶領域を割当てて複数の仮想ボリュームを形成すると共に、ある仮想ボリュームに対するホストプロセッサからの入出力を処理する複数の仮想化装置を有するストレージシステムにおける仮想ボリュームに対するストレージ装置の記憶領域の割り当て方法であって、
複数の仮想化装置に対し、ホストプロセッサから受け付けている処理中の入出力を全て完了させ、以降受け付ける入出力の処理を一時保留させるための要求を発行するステップと、
該要求に対する処理中の入出力の処理の完了報告を、複数の仮想化装置から受けるステップと、
該要求を発した全ての仮想化装置から該完了報告を受けた場合、全ての仮想化装置に対してストレージ装置の記憶領域の割り当て変更の指示を送るステップと、
全ての仮想化装置からの該割り当ての変更の完了報告を受けるステップと、
一時保留にされた該入出力の状態を解除するための指示を全ての仮想化装置に送るステップと、
を有することを特徴とする記憶領域の割り当て方法。

【請求項 2】

前記仮想ボリュームと、該ストレージ装置の実領域となる記憶領域を対応付けする構成情報を格納するテーブルを予めメモリに用意し、
前記記憶領域の割り当て変更の指示を送るに際して、構成情報の差分情報を送信し、該仮想化装置は、該テーブルの該当するエントリの構成情報を変更することを特徴とする請求項 1 記載の記憶領域の割り当て方法。

【請求項 3】

前記入出力処理の完了報告を受取らなかった仮想化装置を制御の対象から外して、記憶領域の割り当て変更を行わないことを特徴とする請求項 1 記載の記憶領域の割り当て方法。

【請求項 4】

該仮想ボリューム上で割り当てが変更される箇所を含むアドレス範囲を対象として、該入出力を一時保留するか否かを制御することを特徴とする請求項 1 記載の記憶領域の割り当て方法。

【請求項 5】

該記憶領域の割り当て変更の指示を送る前に、既に仮想ボリュームが割り当てられたストレージ装置の記憶領域から、新たに仮想ボリュームが割り当てられるストレージ装置の記憶領域へデータをコピーするステップと有することを特徴とする請求項 1 記載の記憶領域の割り当て方法。

【請求項 6】

複数の記憶領域が規定され得るストレージ装置と、該ストレージ装置が持つある記憶領域を割当てて複数の仮想ボリュームを形成すると共に、ある仮想ボリュームに対するホストプロセッサからの入出力を処理する複数の仮想化装置と、該仮想ボリュームに対する該ストレージ装置の記憶領域の割り当ての構成を変更するための構成変更制御装置とを有し、
該構成変更制御装置は、全ての仮想化装置に対して、該入出力の一時保留を要求する手段と、
該要求を受けた該仮想化装置は処理中の入出力を完了すると共に、以降ホストプロセッサからの入出力要求を一時保留する状態にして、該構成変更制御装置に該入出力の処理の完了報告を返す手段と、
該構成変更制御装置は、要求を発した先の全ての仮想化装置からの該完了報告を受けた場合に、該仮想化装置に対して、仮想ボリュームに対するストレージ装置の記憶領域の割り当ての変更を指示する手段と、
を有することを特徴とするストレージシステム。

【請求項 7】

前記構成変更制御装置は、
前記要求手段、該仮想化装置からの完了報告を受領する手段、及び該変更指示手段を有する構成変更制御プログラムと、
該構成変更制御プログラムを実行するプロセッサと、
該仮想ボリュームと該ストレージ装置の実領域となる記憶領域を対応付けする構成情報を登録する構成情報テーブルと、該構成情報の変更前後の差分を記録する差分情報テーブルを格納するメモリとを有し、
前記仮想化装置は、
構成変更の処理を行う構成管理プログラムと、
該構成管理プログラムを実行するプロセッサと、
該仮想ボリュームと該ストレージ装置の実領域となる記憶領域を対応付けする構成情報を登録する構成情報テーブルと、該構成情報の変更前後の差分を記録する差分情報テーブルを格納するメモリを有し、
該構成変更制御装置は、記憶領域の割当て変更の指示を送るに際して、該差分情報テーブルを参照して構成情報の差分情報を該仮想化装置に送信し、該仮想化装置は、該構成管理プログラムを該プロセッサで実行して、受信した差分情報に従って自らの該構成情報テーブルの該当するエントリの構成情報を変更することを特徴とする請求項6記載のストレージシステム。

【請求項8】

更に、該構成変更制御装置に対して構成情報の変更の要求を入力する入力部及び構成変更の状況を表示する表示部を有する管理コンソールを含むことを特徴とする請求項6記載のストレージシステム。

【請求項9】

ストレージ装置が有するある記憶領域を割当てて複数の仮想ボリュームを形成すると共に、ある仮想ボリュームに対するホストプロセッサからの入出力を処理する複数の仮想化装置において、
仮想ボリュームとストレージ装置の実領域となる記憶領域を対応付けの構成を変更するための構成変更制御プログラムと、
該構成変更制御プログラムを実行する第一のプロセッサとを有し、
該プログラムは、
仮想ボリュームとストレージ装置の実領域となる記憶領域を対応付けの構成を変更する前に、入出力の一時保留を他の仮想化装置に対して要求する手段と、
該要求を受けた他の仮想化装置は処理中の入出力の完了すると共に、以降ホストプロセッサからの入出力要求を一時保留する状態に遷移して、完了報告を返す手段と、
他の仮想化装置からの完了報告を受けた場合に、該他の仮想化装置に対して、仮想ボリュームに対するストレージ装置の記憶領域の割り当ての変更を指示する手段と、
他の仮想化装置からの割当ての変更の完了報告を受ける手段と、
一時保留にされた該入出力の状態を解除するための指示を他の仮想化装置に送る手段と、
を有することを特徴とする仮想化装置。

【請求項10】

更に、仮想ボリュームとストレージ装置の実領域となる記憶領域を対応付けする構成情報を登録する構成情報テーブル、及び該構成情報の変更前後の差分を記録する差分情報テーブルを格納するメモリと、
該構成変更制御プログラムをからの要求を受けて、入出力の一時保留、構成情報の変更を行うための構成管理プログラムと、
該構成管理プログラムを実行する第二のプロセッサを有し、
該第二のプロセッサで構成管理プログラムの実行することにより、該構成情報テーブルの内容を更新することを特徴とする請求項9記載の仮想化装置。

【請求項11】

前記第一のプロセッサと該第二のプロセッサとは、同じプロセッサにより構成されること

を特徴とする請求項 10 記載の仮想化装置。

【請求項 12】

前記構成変更制御プログラムは、更に前記構成変更制御プログラムの各手段を実行させる該第一のプロセッサを限定するための調停処理を行う手段を有する請求項 9 記載の仮想化装置。

【請求項 13】

複数面から成る前記構成情報テーブルが用意され、夫々の面のテーブルが切替えられて使用されることを特徴とする請求項 10 記載の仮想化装置。

【請求項 14】

更に、該仮想ボリュームが対応するある記憶領域から他の記憶領域に構成の変更をする場合、該他の記憶領域にデータをコピー処理するためのコピー処理プログラムと、該コピー処理プログラムによるデータのコピー処理の進捗状況を管理するコピー進捗テーブルを有することを特徴とする請求項 9 乃至 13 のいずれかに記載の仮想化装置。

【請求項 15】

実記憶領域を提供するための複数の記憶領域と、該記憶領域を割当てて複数の仮想ボリュームを形成すると共に、ある仮想ボリュームに対するホストプロセッサからの入出力を処理する仮想化装置を有するストレージ装置であって、該仮想化装置は、仮想ボリュームとストレージ装置の実領域となる記憶領域を対応付けの構成を変更する前に、入出力の一時保留を他の仮想化装置に対して要求する手段と、該要求を受けた他の仮想化装置は処理中の入出力の完了すると共に、以降ホストプロセッサからの入出力要求を一時保留する状態に遷移して、完了報告を返す手段と、他の仮想化装置からの完了報告を受けた場合に、該他の仮想化装置に対して、仮想ボリュームに対する記憶領域の割り当ての変更を指示する手段と、他の仮想化装置からの割り当ての変更の完了報告を受ける手段と、一時保留にされた該入出力の状態を解除するための指示を他の仮想化装置に送る手段と、を有することを特徴とするストレージ装置。

【請求項 16】

上記各手段を実現するための構成変更制御プログラムと、該プログラムを実行するプロセッサを有することを特徴とする請求項 15 記載のストレージ装置。

【請求項 17】

該仮想ボリュームに対する割り当ての対象となるある記憶領域から他の記憶領域に構成の変更をする場合、ある記憶領域から他の記憶領域にデータをコピーするためのコピー制御部を有することを特徴とする請求項 15 又は 16 記載のストレージ装置。

【請求項 18】

仮想ボリュームに対するホストプロセッサからの入出力を処理する複数の仮想化装置における仮想ボリュームに対するストレージ装置の記憶領域の割り当てを変更する方法において、複数の仮想化装置に対し、ある時点以降ホストプロセッサから受け付ける入出力の処理を一時保留させるための要求を発し、該要求に対して処理中の入出力の処理が完了した旨の報告を各仮想化装置から受けたことを条件に、各仮想化装置に対してストレージ装置の記憶領域の割り当て変更を行わせ、各仮想化装置からの割り当ての変更の完了報告を受けた後、一時保留させていた該入出力の状態を解除することを特徴とする変更方法。

【請求項 19】

管理コンソールから構成変更の指示を入力するステップを含み、該入力指示に従って前記一時保留させるための要求が発せられることを特徴とする請求項 18 記載の変更方法。

【請求項 20】

ストレージ装置が持つある記憶領域を割当てて複数の仮想ボリュームを形成すると共に、ある仮想ボリュームに対するホストプロセッサからの入出力を処理する複数の仮想化装置を含むストレージシステムにおける仮想ボリュームに対するストレージ装置の記憶領域の

割り当ての変更を行うプログラムにおいて、
複数の仮想化装置に対し、ホストプロセッサからの受け付けている処理中の入出力を完了させ、以降受け付ける入出力の処理を一時保留させるための要求を発行する手段と、
該要求に対する処理中の入出力の処理が完了した旨の報告を、複数の仮想化装置から受ける手段と、
該要求を発した先の全ての仮想化装置から該完了報告を受けた場合、全ての仮想化装置に対してストレージ装置の記憶領域の割り当て変更を指示する手段と、
全ての仮想化装置からの該割り当ての変更の完了報告を受ける手段と、
一時保留にされた該入出力の状態を解除するための指示を全ての仮想化装置に送る手段と、
を有する構成変更のためのプログラム。

【書類名】 明細書**【発明の名称】 ストレージシステム****【技術分野】****【0 0 0 1】**

本発明はストレージシステムに係り、特に冗長構成の仮想化装置を有するストレージシステムにおける仮想ボリュームに対する実ストレージ領域の割り当ての変更に関する。

【背景技術】**【0 0 0 2】**

所謂 SAN (Storage Area Network) において、ストレージ装置の記憶領域に関連して仮想的な複数のボリュームを設定し、ホストプロセッサからネットワークを介してそれらのボリュームを使用するシステムが知られている。

【0 0 0 3】

ネットワークを介して接続されたストレージ装置を仮想化してホストプロセッサからの入出力を可能にするストレージ仮想化装置に関しては、例えば特開 2 0 0 0 - 2 4 2 4 3 4 号公報 (特許文献 1) に開示されているようなシステムがある。この技術によれば、記憶装置システム 1 とホスト 3 0 との間にスイッチ装置 2 0 を設け、このスイッチ装置 2 0 でホスト 3 0 に提供される仮想的な記憶装置システムの仮想化の設定を変更することにより、ホストに提供される仮想的な記憶領域が割り当てられているストレージ装置を変更する。

【0 0 0 4】

【特許文献 1】 特開 2 0 0 0 - 2 4 2 4 3 4 号公報

【発明の開示】**【発明が解決しようとする課題】****【0 0 0 5】**

特許文献 1 に示されるストレージ仮想化装置 (例えばディスクアレイスイッチ) 群の構成情報は、各ストレージ仮想化装置ごとに独立に管理されている。

【0 0 0 6】

このようなストレージ仮想化システムにおいて、システムの稼動中に構成情報を変更することは、入出力の処理中にその宛先が変更されることとなり、データの破壊や入出力障害を招く。そのために、構成変更中に同時に稼動するストレージ仮想化装置を 1 台のみに縮退する方法も考えられるが、負荷の集中や耐障害性が低下するという問題が生じる。上記特許文献 1 には、記憶装置システムが稼動中に、あるボリュームを他のボリュームに移行すると言った構成情報の変更については言及されていない。

【0 0 0 7】

本発明の目的は、冗長構成の仮想化装置を有するストレージシステムにおいて、システムの稼動中にデータ破壊や入出力障害を回避して、ストレージ仮想化装置の構成情報を変更することにある。

【課題を解決するための手段】**【0 0 0 8】**

本発明は、仮想ボリュームに対するホストプロセッサからの入出力を処理する複数の仮想化装置において、複数の仮想化装置に対し、ある時点以降ホストプロセッサから受け付ける入出力の処理を一時保留させるための要求を発し、この要求に対して処理中の入出力の処理が完了した旨の報告を各仮想化装置から受けたことを条件に、各仮想化装置に対してストレージ装置の記憶領域の割り当て変更を行わせ、各仮想化装置からの割り当ての変更の完了報告を受けた後、一時保留させていた入出力の状態を解除するように構成する。

【0 0 0 9】

ストレージシステムに関する好ましい例では、複数の記憶領域が規定され得るストレージ装置と、このストレージ装置が持つある記憶領域を割り当てて複数の仮想ボリュームを形成すると共に、ある仮想ボリュームに対するホストプロセッサからの入出力を処理する複数の仮想化装置とを有するストレージシステムにおいて、仮想ボリュームに対するストレ

ージ装置の記憶領域の割り当ての構成を変更するための構成変更制御装置とを有し、構成変更制御装置は、全ての仮想化装置に対し、構成変更の前に入出力の一時保留を要求する手段と、この要求を受けた全ての仮想化装置は処理中の入出力の完了すると共に、以降ホストプロセッサからの入出力要求を一時保留する状態に遷移して、構成変更制御装置に完了報告を返す手段と、構成変更制御装置は、要求を発した先の複数の仮想化装置からの完了報告を受けた場合に、その仮想化装置に対して、仮想ボリュームに対するストレージ装置の記憶領域の割り当ての変更を指示する手段とを有して構成される。

【0010】

更に、複数の仮想化装置に関する好ましい例としては、仮想ボリュームとストレージ装置の実領域となる記憶領域を対応付けの構成を変更するための構成変更制御プログラムと、構成変更制御プログラムを実行する第一のプロセッサとを有し、この構成変更制御プログラムは、仮想ボリュームとストレージ装置の実領域となる記憶領域を対応付けの構成を変更する前に、入出力の一時保留を他の仮想化装置に対して要求する手段と、要求を受けた他の仮想化装置は処理中の入出力の完了すると共に、以降ホストプロセッサからの入出力要求を一時保留する状態に遷移して、完了報告を返す手段と、他の仮想化装置からの完了報告を受けた場合に、他の仮想化装置に対して、仮想ボリュームに対するストレージ装置の記憶領域の割り当ての変更を指示する手段と、他の仮想化装置からの割り当ての変更の完了報告を受ける手段と、一時保留にされた該入出力の状態を解除するための指示を他の仮想化装置に送る手段とを有して構成される。

【0011】

更にまた、ストレージ装置に関する構成の例として、実記憶領域を提供するための複数の記憶領域と、この記憶領域を割り当てて複数の仮想ボリュームを形成すると共に、ある仮想ボリュームに対するホストプロセッサからの入出力を処理する仮想化装置を有するストレージ装置であって、この仮想化装置は、仮想ボリュームとストレージ装置の実領域となる記憶領域を対応付けの構成を変更する前に、入出力の一時保留を他の仮想化装置に対して要求する手段と、要求を受けた他の仮想化装置は処理中の入出力の完了すると共に、以降ホストプロセッサからの入出力要求を一時保留する状態に遷移して、完了報告を返す手段と、他の仮想化装置からの完了報告を受けた場合に、該他の仮想化装置に対して、仮想ボリュームに対する記憶領域の割り当ての変更を指示する手段と、他の仮想化装置からの割り当ての変更の完了報告を受ける手段と、一時保留にされた該入出力の状態を解除するための指示を他の仮想化装置に送る手段とを有するストレージ装置として構成される。

【0012】

更に好ましい例では、構成情報の変更と同期して、ストレージ装置間でデータのコピーを行うことにより、システムの稼動中に一方のストレージ装置から他方のストレージ装置にデータを移行することができる。また、構成変更制御の機能および上記コピー処理機能をストレージ装置で行うことにより、コピー機能を持たない仮想化装置でもデータ移行を可能にする。

【発明の効果】

【0013】

本発明によれば、システムの稼動中に、入出力の一時保留の影響を極力回避して、仮想ボリュームの構成情報の変更を行うことができる。これより、システムの稼動中に一方のストレージ装置のデータを他方のストレージ装置にデータを移行することができる。さらにコピー機能を持たない仮想化装置でもデータ移行が可能になる。冗長構成で稼動中に、仮想ボリュームの割り当て先を変更可能なストレージ装置を実現できる。

【発明を実施するための最良の形態】

【0014】

以下、本発明の実施の形態を、図面を用いて説明する。
まず、図1から図8を用いて第一の実施形態について説明する。

【0015】

図1は、ストレージシステムの全体の構成を示す図である。

システムは、少なくとも1台のホストプロセッサ12、複数のストレージ装置13、複数の仮想化スイッチ11、構成変更制御装置16及び管理コンソール14が、LANのようなネットワーク15に接続されて構成される。

ホストプロセッサ12は、ストレージ装置13に格納されるデータを使用する計算機である。ホストプロセッサ12は、仮想化スイッチ11が提供する記憶領域を仮想化スイッチ11に接続されていない他の計算機に提供する機能を有するファイルサーバでも良い。

【0016】

ストレージ装置13は、記憶装置13a又は記憶装置システム13bにより構成される。ここで、記憶装置13aは、例えばハードディスクドライブやDVDドライブ等の単体の記憶装置であり、記憶装置システム13bは、複数の記憶装置13a及びこれらを制御する制御装置1301を有するストレージサブシステムである。記憶装置131は、記憶装置13aの記憶領域を論理ユニット（以下「LU」と言う）131として構成する。LU131は、論理的な記憶領域であり、ホストプロセッサ12などストレージ装置13に接続される装置には、論理的に独立した1つのストレージ装置として認識される。

【0017】

また、論理ユニット131は、複数の部分的な論理的記憶領域（以下「実領域」）132から構成される。各実領域132も、ストレージ装置13が有する物理的な記憶領域に対応する。実領域132のサイズは任意であり、その範囲は、連続したアドレスをもつ領域である。

【0018】

仮想化スイッチ11は、図示するように通信線やスイッチによって他の装置と接続されており、他の装置と通信することができる装置である。また、仮想化スイッチ11は、仮想化スイッチ11自身に接続されている複数のストレージ装置13が有する記憶領域をまとめて一つまたは複数の記憶領域とする（所謂、仮想化する）仮想化装置である。そして、仮想化スイッチ11は、仮想化した記憶領域を、仮想化スイッチ11に接続されるホストプロセッサ12に提供する。仮想化スイッチ11がホストプロセッサ12に対して提供する仮想的な記憶領域を以下、仮想ボリューム100と称する。

【0019】

仮想化スイッチ11とホストプロセッサ12との間および仮想化スイッチ11とストレージ装置13との間で使用される通信線やスイッチでは、ファイバチャネル等のプロトコルが使用される。ここで、使用される通信線やプロトコルはローカルエリアネットワーク等で使用される通信線やプロトコルでも良い。仮想化スイッチ11は、ホストプロセッサ12とストレージ装置13との間に接続され、ホストプロセッサ12が発行するコマンドをストレージ装置13側に転送する機能を有する。

【0020】

仮想ボリューム100は、少なくとも1つの実領域132から構成される仮想化された記憶領域である。仮想化スイッチ11は、少なくとも1つの仮想ボリューム100をホストプロセッサ12に提供することができる。各々の仮想ボリューム100には、仮想ボリュームを特定するための仮想化スイッチ11内で一意の識別子（以下「仮想ボリューム識別子」）が付与される。また、個々の仮想ボリューム100の記憶領域には連続したアドレスが付されている。ホストプロセッサ12は、ストレージ装置13のLU131内の実領域132を直接指定しないで、仮想ボリューム識別子及び仮想ボリューム100内の場所を示すアドレスを指定して、ストレージ装置13に格納されたデータをアクセスする。

【0021】

管理コンソール14は、仮想ボリューム100を作成するためにシステム管理者によって使用されるパーソナルコンピュータPCのような計算機であり、表示装置や入力装置、及びメモリ等を備える。管理コンソール14は、LAN15を介して仮想化スイッチ11と接続され、互いに通信することができる。

【0022】

構成変更制御装置16は、仮想化スイッチ11を制御して構成情報、即ち仮想ボリュー

ム 100 と実領域 132 との対応付けの変更を制御する装置である。構成変更制御装置 16 は、例えば、PC やサーバによって構成することができ、LAN 15 を介して仮想化スイッチ 11 と接続され、互いに通信することができる。

【0023】

図 2 は、図 1 の仮想化スイッチ 11 の内部構成を示す。

仮想化スイッチ 11 は、入力ポート 240、出力ポート 250、転送部 230、プロセッサ 210、メモリ 220、バス 270 および通信部 260 から構成される。転送部 230、プロセッサ 210、メモリ 220 および通信部 260 は全てバス 270 に接続され、互いにデータを送受信する。

【0024】

入力ポート 240 は、仮想化スイッチ 11 がホストプロセッサ 12 と通信するための通信線を接続するポートである。出力ポート 250 は、仮想化スイッチ 11 がストレージ装置 13 と通信するための通信線を接続するポートである。尚、入力ポート 240 及び出力ポート 250 を構成する要素は、共通の場所にあっても良い。この場合、どのポートを入力ポートあるいは出力ポートとして使用するかは、使用者が選択する。

【0025】

転送部 230 は内部メモリを有し、そのメモリに転送情報テーブル 231 を保持する。転送情報テーブル 231 には、各入力ポート 240 を介して仮想化スイッチ 11 と通信可能なホストプロセッサ 12、及び各出力ポート 250 を介して仮想化スイッチ 11 と通信可能なストレージ装置 13 との間の対応関係についての情報が格納される。

【0026】

転送部 230 は、転送情報テーブル 231 を参照し、入力ポート 240 でホストプロセッサ 12 から受信した入出力要求を、要求先のストレージ装置 13 と仮想化スイッチ 11 との間の通信に使用される出力ポート 250 へ転送する。また、転送部 230 は、出力ポート 250 でストレージ装置 13 から受信した応答情報やデータを、受信すべきホストプロセッサ 12 と仮想化スイッチ 11 の間の通信に使用される入力ポート 240 へ転送する。ただし、ホストプロセッサ 12 から受け取った入出力要求が仮想ボリューム 100 に対する入出力要求である場合には、転送部 230 は、後述する入力キュー 241 にその入出力要求をエンキューし、プロセッサ 210 にその処理を依頼する。また、後述する出力キュー 251 に積まれた入出力要求を、出力ポートを介してストレージ装置 13 へ転送する。

【0027】

プロセッサ 210 はメモリ 220 上に格納されたプログラムを実行し、ホストプロセッサ 12 から仮想ボリューム 100 宛ての入出力の処理や構成情報の変更処理を行う。

【0028】

メモリ 220 は、プロセッサ 210 が実行するプログラムやその実行に必要な情報を格納する。メモリ 220 が格納するプログラムおよびデータには、構成管理プログラム 211、I/O 処理プログラム 213、入力キュー 241、出力キュー 251、保留中キュー 242、処理中キュー 252、構成情報 221 および構成情報差分 222 が含まれる。

【0029】

構成管理プログラム 211 は、構成変更制御装置 16 からの要求を受けて、入出力の一時保留および再開処理や構成情報の変更処理を行うためのプログラムである。

I/O 処理プログラム 213 は、ホストプロセッサ 12 からの仮想ボリューム 100 宛ての入出力の処理、即ち入出力をストレージ装置 13 宛ての入出力へ変換して転送する処理のためのプログラムである。

【0030】

入力キュー 241 は転送部 230 が仮想ボリューム 100 宛ての入出力要求を積むためのキューである。出力キュー 251 は I/O 処理プログラム 213 が処理したストレージ装置宛ての入出力要求を積むためのキューである。入力キュー 241 および出力キュー 251 の数は任意である。

【0031】

保留中キュー 242 は、仮想化スイッチ 11 が入出力の処理を一時的に保留している状態（I/O一時保留状態）のときに受け付けた仮想ボリューム 100 宛ての入出力要求を記憶しておくためのキューである。処理中キュー 252 は、仮想スイッチ 11 が処理してストレージ装置 13 宛てに転送した仮想ボリューム 100 宛ての入出力要求を、該入出力が完了するまで記憶しておくためのキューである。

【0032】

構成情報 221 とは、仮想ボリューム 100 と実領域 132 との対応付けの表である。構成情報差分 222 とは、構成情報 221 を変更する際の変更前後の差分を記録するバッファである。

【0033】

なお本実施形態では、保留中キュー 242、処理中キュー 252、構成情報 221 および構成情報差分 222 は仮想化スイッチ 11 ごとに 1 個ずつである。通信部 260 はプロセッサ 210 が LAN 15 を介して、構成変更制御装置 16 および管理コンソール 14 と通信可能にする。

【0034】

図 3 は、図 1 の構成変更制御装置 16 の内部構成を示す。構成変更制御装置 16 は、例えば管理サーバのような装置であり、プロセッサ 161、メモリ 162、バス 270 および通信部 260 から構成される。プロセッサ 161、メモリ 162 および通信部 260 は全てバス 270 に接続され、互いにデータを送受信する。プロセッサ 161 はメモリ 162 上に格納されたプログラムを実行し、仮想化スイッチ 11 の構成変更を制御する。

【0035】

メモリ 162 は、プロセッサ 161 が実行するプログラムやその実行に必要な情報を格納する。メモリ 162 が格納するプログラムおよびデータには、構成変更制御プログラム 212、構成情報 221 および構成情報差分 222 が含まれる。構成変更制御プログラム 212 は、仮想化スイッチ 11 の構成変更を制御するためのプログラムである。

【0036】

図 4 は、ストレージシステムにおける構成情報 221 のテーブルを示す図である。構成情報のテーブル 221 は、仮想化スイッチ 11 及び構成変更制御装置 16 内に備えられる。

構成情報 221 は、仮想ボリュームのアドレス 41、オフセット 42、サイズ 43、LU のアドレス 44、及びオフセット 45 を含むエントリにより構成される。各エントリは、実領域 132 および実領域 132 が割当てられている仮想ボリューム 100 上の部分領域に対応付けられている。LU のアドレス 4 はエントリと対応する実領域 132 が含まれる LU 131 を仮想化スイッチ 11 が識別するための情報を示し、オフセット 45 は実領域 132 の LU 131 上の開始アドレスを示し、サイズ 43 は実領域 132 のサイズを示す。

【0037】

仮想ボリュームのアドレス 41 は、ホストプロセッサ 12 が仮想ボリューム 100 を識別するための情報、オフセット 42 は、そのエントリに対応する部分領域の仮想ボリューム 100 上の開始アドレスを示す。なお、仮想ボリュームのアドレス 41 および LU のアドレス 44 は、具体的にはファイバチャネルの WWN (World Wide Name) またはポート ID、および LUN (logical Unit Number) との組を使用する。

【0038】

次に、図 5 を参照して、このストレージシステムにおける構成変更について概略説明する。図 5 は、ストレージシステムにおける構成変更制御装置 16 と仮想化スイッチ 11 との通信手順を示す図である。

【0039】

管理コンソール 14 から構成情報の変更要求の指示が入力されると、それは構成変更制御装置 16 に伝えられる (501)。そしてステップ 502 から 505 の処理を実行して、ホストプロセッサ 12 が発行した入出力の処理を一時保留する。これは構成変更の前後で処理中の入出力が存在すると矛盾が生じるためである。もしこのように入出力処理中に構成情報が変更されると、その入出力は正しい実領域 132 と対応付けが出来なくなり、データ破壊を生じる可能性がある。このデータ破壊は以下の処理により防止される。

【0040】

すなわち、構成変更の指示を受けた構成変更制御装置 16 は、全仮想化スイッチ 11 に対して入出力処理を一時保留する要求 (IO一時保留要求) を発行する (502)。この保留要求を受けた仮想化スイッチ 11 は、処理中の入出力及び以後の入出力処理を一時保留して、処理中の入出力の実行完了を待つ (503)。実行完了後、構成変更制御装置 16 に対し、ステップ 503 の完了の報告 (IO一時保留完了報告) をする (504)。その後、構成変更制御装置 16 は全ての仮想化スイッチ 11 からの完了報告を待つ (505)。

【0041】

以上の処理で構成情報の変更が可能になった。そして次に、ステップ 506 から 509 が実行され、全ての仮想化スイッチ 11 の構成情報を一括して変更する。これは仮想化スイッチ 11 が持つ構成情報の一貫性が取れていないと、処理する仮想化スイッチ 11 により、同じ仮想ボリュームに対応する入出力が異なる実領域 132 と対応付けられることでデータ破壊を生じる可能性があるためである。

【0042】

構成変更制御装置 16 は、全ての仮想化スイッチ 11 に構成情報の差分とその差分を用いた構成情報の変更要求を送信する (506)。要求を受けた仮想化スイッチ 11 は、構成情報を変更する (507)。そして構成変更が完了した後、構成変更制御装置 16 に対し、構成情報変更の完了を報告する (508)。その後構成変更制御装置 16 は全ての仮想化スイッチ 11 からの完了報告を待つ (509)。

【0043】

最後にステップ 510 から 513 を実行し、ステップ 503 で一時保留した入出力を再開する。

構成変更制御装置 16 は、全ての仮想化スイッチ 11 に対し一時保留している入出力の処理を再開する要求 (IO再開要求) を発行する (510)。要求を受けた仮想化スイッチ 11 は、一時保留中の入出力を再開する (511)。そして構成変更制御装置 16 に対し、再開処理の完了を報告する (512)。その後、構成変更制御装置 16 は全ての仮想化スイッチ 11 からの完了報告を待ち (513)、管理コンソールへ完了報告を行う (514)。

このようにして、入出力処理中 (即ちシステムの稼動中) の構成情報の矛盾や複数の仮想化スイッチ 11 の間での構成情報の矛盾を発生させずに、構成変更が可能になる。

【0044】

次に、図 6 を参照して、構成変更制御装置 16 における構成変更制御プログラム 212 の処理動作について説明する。

【0045】

まず、プロセッサ 161 は、全ての仮想化スイッチ 11 に対し、IO一時保留要求を発行し、全ての仮想化スイッチ 11 からの完了報告を待つ (601)。待ち合わせが正常に完了した場合 (602 “Y”)、構成情報 221 の変更が可能になる。この場合、プロセッサ 161 は全ての仮想化スイッチ 11 に対し、構成変更要求と構成情報差分 222 を送信して構成情報 221 を変更させる。そして全ての仮想化スイッチ 11 からの完了報告を待つ (603)。待ち合わせが正常に完了した場合 (604 “Y”)、構成情報 221 が正常に変更されたので、保留中の入出力を再開するために、プロセッサ 161 は全ての仮想化スイッチ 11 に対し、IO再開要求を送信し、全ての仮想化スイッチ 11 からの完了報告を待つ (605)。待ち合わせが正常に完了した場合 (606 “Y”)、全ての変更

処理が成功したと判断して、管理コンソール 14 に成功した旨を示すメッセージを返す（607）。以上が構成変更制御プログラム 212 の正常処理の流れである。

【0046】

一方、仮想化スイッチ 11 が I/O 一時保留要求を正常に完了しなかった場合（602 “N”）、以後の処理を続行するか否かを判断する（608）。本実施形態では、成功の応答を返した仮想化スイッチ 11 が 2 台以上ある場合、処理を続行し（608 “Y”）、成功の応答を返さなかった仮想化スイッチ 11 は使用しないようにするために、制御対象から外し（611）、ステップ 603 から処理を継続する。なおステップ 611 で、仮想化スイッチ 11 のいくつかを制御対象から外した場合、これらの仮想化スイッチ 11 の入出力を誤って再開してしまうと不正な構成情報 221 で運用されてしまうことになるので、プロセッサ 161 はステップ 607 で、管理コンソール 14 に対して制御対象から外した仮想化スイッチ 11 を報告し、「構成変更の際に入出力の一時保留の同期に失敗して構成を変更できなかったため仮想化スイッチのいくつかを停止した」旨のメッセージを管理コンソール 14 の表示装置に表示する。

【0047】

また、処理を続行しない場合（608 “N”）、プロセッサ 161 はまず全ての仮想化スイッチ 11 に対して I/O 再開要求を送信し、全ての仮想化スイッチ 11 からの完了報告を待ち（609）、ステップ 601 以降を再試行するか否かを判定する（610）。ここで、ステップ 602 における異常がタイムアウトであった場合、このタイムアウトは仮想化スイッチ 11 が多数の入出力を処理中のためにそれらの完了を待ちきれないことに起因すると考えられる。従って、再試行によりこの完了待ちが成功する可能性があるため、他に異常がない場合は、本実施形態では一定回数に限り再試行する（610 “Y”）。ステップ 610 “Y” 以降、プロセッサ 161 はステップ 601 から処理を繰り返す。

【0048】

一方、再試行しない場合（610 “N”）には、ステップ 609 で成功が返らなかった仮想化スイッチ 11 が有れば、これを制御対象から外す（617）。そして、管理コンソール 14 に対してエラーの発生とその要因を報告し、「仮想化スイッチの入出力完了待ちに失敗した」旨のメッセージを表示する（616）。より望ましくは、管理コンソール 14 は、「ホストプロセッサ 12 から仮想ボリューム 100 への入出力頻度が低いときに構成変更の再試行せよ」と、管理者に促す旨のメッセージを表示する。

【0049】

さらに、ステップ 604 で異常検出、即ち構成情報 221 の変更失敗した場合は、プロセッサ 161 は元の構成情報 221 に戻して（612、613）入出力を再開させた（614、615）後に、管理コンソール 14 に対してエラーの発生とその要因を報告し、例えば「管理者が指定した構成情報への変更に失敗した」という旨のメッセージを表示する（616）。

【0050】

なお、ステップ 613 および 615 で異常を検出した場合（“N” の場合）は、以降の処理を続行できないので、プロセッサ 161 は成功の応答が返らなかった仮想化スイッチ 11 を制御の対象から外し（617）、ステップ 616 の処理を行い、例えば「管理者が指定した構成情報への変更に失敗しリカバリ処理にも失敗した」旨のメッセージを表示する。以上の様に、プロセッサ 161 は構成情報 221 の変更を制御する。

【0051】

次に、図 7 を参照して、仮想化スイッチ 11 における構成管理プログラム 211 の処理動作について説明する。

構成管理プログラム 211 が呼び出される要因は、仮想化スイッチ 11 が構成変更制御装置 16 から、I/O 一時保留要求、I/O 再開要求、構成情報変更要求のいずれかを受領した場合である。まず、I/O 一時保留要求を受領した場合（701 “Y”）、プロセッサ 210 は、I/O 処理プログラム 213 を呼び出して I/O 一時保留状態にする（702）。そして一時保留状態への遷移に成功した場合（703 “Y”）、プロセッサ 210 は、構成変

更プログラム 212 を実行するプロセッサ 161 に対して成功の応答を返す (704)。

【0052】

次に、構成情報変更要求を受領した場合 (705 “Y”)、プロセッサ 210 は、I/O 一時保留状態であることを確認する (706)。これは誤って入出力の処理中に構成情報 221 を変更しないためである。さらにプロセッサ 210 は構成情報 221 を変更する (707)。その結果、正常に変更できた場合には (708 “Y”)、ステップ 704 を実行する。

【0053】

また、I/O 再開要求を受領した場合 (709 “Y”)、プロセッサ 210 は、I/O 一時保留状態であることを確認する (710)。その結果、保留状態であれば (710 “Y”)、I/O 処理プログラム 213 を呼び出して I/O 一時保留状態を解除して入出力を再開する (711)。正常に変更できた場合 (712 “Y”)、ステップ 704 を実行する。以上が構成管理プログラム 211 の正常パスである。

【0054】

一方、異常パスは、ステップ 703、706、708、710、712 で、否 (“N”) になった場合であり、いずれの場合もエラーとその要因をプロセッサ 161 に返す (713)。

このように、プロセッサ 210 は構成管理プログラム 211 を実行することで、構成変更制御装置 16 の指示に従って構成情報 221 を変更できる。

【0055】

次に、図 8 を参照して、仮想化スイッチ 11 における I/O 処理プログラム 213 の処理動作について説明する。

I/O 処理プログラム 213 が呼び出される要因は、入力キュー 241 に入出力要求がエンキューされたとき、処理中キュー 252 の入出力要求が完了したとき、構成管理プログラム 211 の処理により、I/O 一時保留状態に遷移するときおよび I/O 再開要求により、I/O 一時保留状態を解除するときである。

【0056】

入力キュー 241 に入出力要求がエンキューされたとき (801) は、プロセッサ 210 は、I/O 一時保留状態でない場合 (802 “N”)、その入出力要求を処理する。即ち構成情報 221 を参照して仮想ボリューム 100 のアドレスから LU131 のアドレスへの変換を行って、出力キュー 251 にエンキュー (以下、入出力要求を処理したと言う) した後、その入出力要求を処理中キュー 252 にエンキューする (803)。その後プロセッサ 210 はイベント待ち (804) を実行し、次の起動要因を待つ。

一方、I/O 一時保留状態にある場合 (802 “Y”)、プロセッサ 210 はその入出力要求を保留中キュー 242 にエンキューし (805)、ステップ 804 以降の処理を実行する。

【0057】

処理中キュー 252 の入出力要求が完了したとき (807 “Y”)、プロセッサ 210 は、処理中キュー 252 から入出力要求をデキューする (808)。そして I/O 一時保留状態か否かを確認する (809)。もし、I/O 一時保留状態でない場合は (809 “N”)、プロセッサ 210 はステップ 804 以降の処理を実行する。

一方、I/O 一時保留状態である場合は (809 “Y”)、プロセッサ 210 は処理中キュー 252 を検査する (810)。検査の結果、空になっていた場合 (810 “Y”) には、構成管理プログラム 211 に I/O 保留状態への遷移完了を報告し (811)、処理を完了する。これに対して処理中キュー 252 が空でない場合 (810 “N”) には、プロセッサ 210 はステップ 804 からの処理を繰り返す。

【0058】

I/O 一時保留状態に遷移するとき (812 “Y”) は、プロセッサ 210 は I/O 一時保留状態になったことを記憶し (813, 814)、ステップ 810 から処理を継続する。

【0059】

また、I/O再開要求を受けた場合には(815 “Y”)、I/O一時保留状態にあるか否かを確認する(816)。確認の結果、保留状態にあれば、I/O一時保留状態を解除し、保留中キュー242の入出力要求を処理して、処理中キュー252へエンキューする(817)。そして正常にエンキューできた場合には(818 “Y”)、ステップ811の処理(即ち完了報告)を実行する。以上が、I/O処理プログラム213が正常に実行された場合の処理である。

【0060】

これに対して、I/O処理プログラム213の処理が正常でない場合には、ステップ814、816、818で否(“N”)となった場合である。これらの場合には、何れも処理を継続できないので構成管理プログラム211にエラーとその要因を返す(819)。

【0061】

この実施形態においては、構成変更制御プログラム212の実行時のステップ610によりデータ破壊の原因となる構成情報221の変更中の入出力処理を回避し、さらにステップ603により同様にデータ破壊の原因となる仮想化スイッチ11間の構成情報221の矛盾を回避する。

【0062】

次に、図9から図11を用いて第二の実施形態について説明する。

図9は、ストレージシステムにおける仮想化スイッチ11の内部構成を示す。

図2に示す仮想化スイッチ11と相違する点は、バス270に接続されたプロセッサ161を備え、メモリ220に構成変更制御プログラム212が格納され、この構成変更制御プログラム212をプロセッサ161が実行することである。ここでプロセッサ161はタイマ165を備え、このタイマ165によりプロセッサ161で構成変更制御プログラム212の処理が周期的に起動される。なお、タイマ165はソフトウェアで実現したタイマであっても良い。

【0063】

本実施形態によれば、仮想化スイッチ11に構成変更制御プログラム212及びプロセッサ161を内蔵することにより、仮想化スイッチ11を構成変更制御装置16として機能させることができる(以下、このように機能付けられ仮想化スイッチ11も構成変更制御装置16と呼ぶ)。仮想化スイッチ11が複数在る場合、それら全てに上記のような構成を備えて、構成変更制御装置16としての機能付けを行うことができる。この様に全ての仮想化スイッチ11に構成変更制御プログラム212及びプロセッサ161が備えられれば、図1に示すシステムにおける構成変更制御装置16は不要となる。

なお、仮想化スイッチ11に設けられるプロセッサ210に、プロセッサ161の機能を兼用させることもできる。

【0064】

図10は、第二の実施形態における構成変更制御プログラム212の処理を示す。図10の処理動作で、調停処理600を含む点が、図6と相違する。その調停処理600の詳細な処理フローを図11に示す。

調停処理600は、複数の仮想化スイッチ11に対応する複数のプロセッサ161のうち、ステップ601以降の処理を実行するプロセッサ(以下、マスタ権を持つプロセッサという)を1個に限定させる役目をする。本実施形態では、構成変更制御プログラム212は、管理コンソール14からの構成変更の要求以外に、他の構成変更制御装置(仮想化スイッチ11に内臓の)16からの監視パケットの受信およびタイマ165によっても起動される。

【0065】

図11のフローチャートにおいて、まず、プロセッサ161は、監視パケットを受信した場合(1101 “Y”)、この監視パケットを送信した構成変更制御装置16へ応答パケットを返し(1102)、調停処理を完了し、構成変更制御プログラム212の処理も完了する(1110)。

【0066】

監視パケットおよび応答パケットは、送信した構成変更制御装置 16 の ID、マスタ権を持つプロセッサを内蔵する構成変更制御装置 16 の ID（以下、マスタ権を持つ装置の ID）を含む。これらの ID は、構成変更制御装置 16 を識別するための識別子であり、例えば通信部 260 が持つ IP アドレスである。プロセッサ 161 が管理コンソールから構成変更の要求を受けた場合（1103 “Y”）、プロセッサ 161 は他の構成変更制御装置 16 にマスタ権を主張する、即ち、自装置 16 の ID を、マスタ権を持つ装置の ID として記録した監視パケットを全ての構成変更制御装置 16 に送信して応答を待つ（1104）。正常終了、即ち自装置 16 以外の ID を、マスタ権を持つ装置の ID として持つ応答パケットを受信しなかった場合（1105 “Y”）、処理を完了して構成変更制御プログラム 212 のステップ 601 に移る（1106）。

【0067】

ステップ 1103 において、タイマ 165 の契機で呼び出された場合、プロセッサ 161 は他の構成変更制御装置 16 を監視する、即ち監視パケットを、マスタ権を持つ装置に送信する（1107）。正常終了、即ちマスタ権を持つ装置からその装置の ID を、マスタ権を持つ装置の ID として持つ応答パケットを受信した場合（1108 “Y”）、ステップ 1110 へ進み、処理を完了する。以上が調停処理の正常な場合である。

【0068】

これに対して、異常な処理の場合、即ちステップ 1105、1108 で否（“N”）の場合には、マスタ権を持つ装置が他にあるため、管理コンソール 14 にその旨を報告して（1109）、ステップ 1110 へ進み、処理を完了する。
その他の処理動作は、前述した第一の実施形態と同様なので、説明を省略する。

【0069】

このように、第二の実施形態によれば、ステップ 1104、1102 の処理動作により、マスタ権を持つ装置が、最初に管理コンソール 14 から構成変更の要求を受けた装置 16 のみに限定される。即ち、システム中で複数のプロセッサ 161 が構成変更制御プログラム 212 を実行しても矛盾を生じさせずに構成情報 221 を変更することができる。なお、ステップ 1107、1102 により、システムの管理者はマスタ権を持つ装置 16 に障害が発生した場合に、管理コンソール 14 を通じてこれを知ることができる。

【0070】

次に、図 12 から図 18 を用いて第三の実施形態を説明する。

図 12 は、ストレージシステムにおける仮想化スイッチ 11 の内部構成を示す。
図 9 に示す仮想化スイッチ 11 と比べて、一時保留制御表 223 を備える点、複数の処理中キュー 252 を持つ点、及び 2 面の構成情報 221a、221b を持つ点が相違する。尚、プロセッサ 210 はタイマ 215 を備える。

【0071】

一時保留制御表 223 は、構成情報 221 の各エントリと処理中キュー 252 との対応付けを記憶することで、入出力を一時保留するか否かを制御する。複数の処理中キュー 252 を持つ理由は、入出力の宛先のアドレスにより処理中キュー 252 を分けることにより、完了待ちおよび一時保留する入出力を限定するためである。構成情報 221 を二面持つのは、構成情報 221 の面の切替えを行うことにより、IO 一時保留状態での処理時間を短縮するためである。また、タイマ 215 は IO 処理プログラム 213 の処理でプロセッサ 210 がセットし、IO 一時保留状態のタイムアウトを検出するために設けられる。

【0072】

図 13 は、第三の実施形態における、構成変更制御装置 16 と仮想化スイッチ 11 との通信手順を示す。

【0073】

この実施形態では、複数の仮想化スイッチ 11 の中で構成変更制御の機能が有効となる仮想化スイッチは 1 台のみであるので、構成変更制御機能が有効である仮想化スイッチ 11 と、その機能が有効でない仮想化スイッチ 11 との間で通信が必要となる。その通信の手順が図 13 に示す例である。また、同じ仮想化スイッチ 11 の中でも同様の通信が行わ

れる。そこで、図 13 に示す「構成変更制御機能」は、図 12 のプロセッサ 161 又はそこで実行される構成変更制御プログラム 212 を指し、「仮想化スイッチ 11」は自他の仮想化スイッチ 11 におけるプロセッサ 210 又はそこで実行される構成管理プログラム 211 を指すもの想定して参照願いたい。

【0074】

図 13 について図 5 に示す通信手順との相違点は、ステップ 501 と 502 の間に、ステップ 531 ～ 533 が追加されている。また、図 5 のステップ 506 ～ 508 で、構成情報変更要求を送信することに代わり、構成情報切替要求を送信する。

【0075】

さて、ステップ 530 で、プロセッサ 161 は全ての仮想化スイッチ 11 に構成情報差分 222 を送信する (530)。全仮想化スイッチ 11 では構成情報の差分を格納し、構成情報 221 を作成する (531)。仮想化スイッチ 11 は構成情報 221 のうち、IO 処理プログラム 213 で使用していない面 (ここでは 221b とする) を作成する。その後、仮想化スイッチ 11 は構成情報差分受信完了報告を構成変更制御装置 16 へ返す (532, 533)。

またステップ 507' で、仮想化スイッチ 11 は構成情報 221a から構成情報 221b への切替を行う。このように IO 一時保留状態の間 (ステップ 503 から 511 の間) での処理が、ステップ 507' による構成情報 221 の面切替えだけで済むようになるため、処理時間を短縮できる。

【0076】

図 14 は、図 12 における一時保留制御表 223 の一例を示す。

一時保留制御表 223 は複数のエントリからなり、それぞれのエントリは一時保留の状態 2231 および処理中キュー ID の 2232 を持つ。一時保留の状態 2231 と処理中キューの ID 2232 は、共に完了待ちおよび一時保留する入出力を限定するために備えられる。なお、一時保留の状態 2231 および処理中キューの ID 2232 の初期値は、仮想化スイッチ 11 の初期化時または仮想ボリューム 100 の作成時にプロセッサ 210 が、それぞれ一時保留でない状態および空いている (一時保留制御表 223 と対応付けられていない) 処理中キュー 252 の ID を格納することにより行う。なお、空いている処理中キュー 252 が無い場合には、プロセッサ 210 が新規に処理中キュー 252 を作成し、その ID を ID 2232 に格納する。

【0077】

図 15 は、図 12 における構成情報 221 を示す図である。

図 15 は、図 4 に示すテーブルに比べ、制御表 223 のインデクス 411 を備えたことが相違する。インデクス 411 は一時保留制御表 223 のエントリを指している。構成情報 221 のエントリ即ち仮想ボリューム 100 上のアドレス範囲が、一時保留制御表 223 のエントリを通して、一時保留の状態 2231 と処理中キュー 252 と対応付けられていることにより、アドレス範囲ごとに入出力を一時保留するか否かを制御することができる。

【0078】

図 16 は、図 12 に示す仮想化スイッチ 11 における構成変更制御プログラム 212 の処理を示す。

図 10 に比べて相違する点は、調停処理 (ステップ 600) の後に、ステップ 625 及び 626 が挿入され、ステップ 602 の後にステップ 627 が挿入されている。また、ステップ 603 及び 612 の代わりにステップ 603' 及び 612' を含んでいる。

【0079】

ステップ 625 とステップ 626 の処理は、図 13 の手順 530、533 と同じである。ステップ 627 において、プロセッサ 161 は、プロセッサ 210 が IO 処理プログラム 213 の処理中に IO 一時保留状態のタイムアウトを検出した場合 (627 “Y”) に、一旦 IO 再開要求を出し (609)、ステップ 601 以降を再試行させる。

【0080】

このステップ627により、ホストプロセッサ12による入出力のタイムアウトに先んじてIO一時保留状態を解除することができる。またステップ603'および612'は構成情報221を書き換える代わりに、構成情報を221aから221bへ切り替える処理を仮想化スイッチ11に依頼する。構成情報221の切り替え処理は、メモリに構成情報を書き換えるよりも短時間で済むため、結果IO一時保留状態の時間を短縮する事が可能になる。

【0081】

図17は、図12の仮想化スイッチ11における構成管理プログラム211の処理を示す。

図7に示すフローチャートとの相違する点は、ステップ709“N”の後にステップ714から718が加わったことである。ステップ707では構成情報221のうち、IO処理プログラム213が使っていない面(221bとする)の更新を行っている。ステップ714“Y”から704までの処理は、図13の手順507'、508'と同じである。

【0082】

構成管理プログラム211が呼び出される要因は、IO一時保留要求・再開要求、構成情報更新要求・構成情報切替要求の他に、IO処理プログラム213からの要求があり、この要求がある場合ステップ601“N”へ進む。次にプロセッサ210は、IO処理プログラム213の要求を構成変更制御装置16へ伝える(718)。この処理により、IO一時停止状態のタイムアウトを契機に、構成変更制御プログラム212での処理を再試行することができる。

さらにステップ705からステップ704の処理は、図13の手順531、532と同じである。

【0083】

図18は、図12におけるIO処理プログラム213の処理を示す。

図18は図8のフローチャートに比べて、ステップ802、803、805、808、809、810、813、816、817の処理内容が相違する。

【0084】

ステップ802'、809'、816'では、プロセッサ210はIO一時保留状態であるか否かの判定に、構成情報221中のインデクス411が指すエントリの一時保留の状態2231を検査することで行う。これにより、入出力の宛先となっている仮想ボリューム100のアドレス範囲ごとに入出力を一時保留するか否かを制御することができる。

【0085】

さらに、ステップ803'および808'では、プロセッサ210は処理中キュー252として、構成情報221中のインデクス411が指すエントリの処理中キューのID2232で特定される処理中キュー252を使用する。

【0086】

また、ステップ813'において、プロセッサ210は、構成情報差分222の内容により定義されている仮想ボリューム100上の領域に対する入出力に限り、IO一時保留状態になるように状態を変更する。具体的には、プロセッサ210は構成情報差分222に登録されている全エントリに対応する、構成情報221aのインデクス411をリストアップし、これらのインデクス411に対応するエントリの一時保留の状態2231を全て一時保留中に変更する。既に述べたように、ステップ802'ではこの一時保留の状態2231を検査している。従って、この一時保留の状態2231に対応する入出力要求、即ち構成情報差分222により影響を受ける入出力要求のみが保留されることになる。

【0087】

ステップ810'において、プロセッサ210は、構成情報差分222に登録されている全エントリに対応する、構成情報221aのインデクス411をリストアップし、これらのインデクス411に対応するエントリの処理中キューのID2232が指す処理中キュー252を全て検査し、空になっているか否かを判定する。

【0088】

ステップ 805' において、プロセッサ 210 は、処理中キュー ID 2232 に示された処理中キュー 252 にエンキューし、タイマ 215 をセットし、ステップ 817' において、プロセッサ 210 はタイマ 215 をリセットして、ステップ 813' で保留中にした一時保留の状態 2231 を一時保留でない状態にリセットし、保留中キュー 242 にエンキューされている入出力を処理し、処理中キュー ID 2232 に示された処理中キュー 252 にエンキューする。

【0089】

ステップ 820 “Y” では、プロセッサ 210 が、タイマ 215 のタイムアウトを契機に IO 処理プログラム 213 の処理を開始する場合である。この場合、プロセッサ 210 はステップ 819 を実行して、構成管理プログラムへエラーを返し、プロセッサ 210 はステップ 818 を処理する。

その他は第二の実施形態と同様なので説明を省略する。

【0090】

このように第三の実施形態では、ステップ 625 と 60' により、IO 一時保留状態の時間を短縮することができ、ステップ 805'、820、627 により、IO 一時保留状態を解除し再試行することで IO 一時保留状態の時間を限定できる。これにより、ホストプロセッサ 12 における入出力のタイムアウトまたは性能低下を防止できる。

【0091】

さらに本実施形態では、ステップ 813' と 802' により一時保留する入出力を構成変更の影響を受ける範囲に限定することができる。これにより構成変更による性能低下を防止できる。

【0092】

なお、本実施形態の変形例として、上記した 2 面の構成情報 221 を持たず、ステップ 625 を実行しない方法もある。例えば、ステップ 603' 即ち構成情報 221 を切り替える処理の代わりにステップ 603 の処理（第一の実施形態における構成情報 221 の書き換え）を行っても良い。その場合、IO 一時保留状態の時間は長くなるが、メモリ 220 の容量を小さく抑えることができる。

【0093】

次に、図 19 から図 26 を用いて第四の実施形態を説明する。

図 19 は、第四の実施形態における仮想化スイッチ 11 の内部構成を示す。

図 12 に示す構成と比べて、コピー処理プログラム 214 とコピー進捗表 224 が追加され、1 面から成る構成情報 221 を持つ点が異なる。

【0094】

コピー処理プログラム 214 は、プロセッサ 210 で実行され、仮想ボリューム 100 が対応する LU 1311 を別の LU 1312 へ変更するときに、LU 1311 から LU 1312 へデータをコピーする処理を行う。コピー進捗表 224 は、コピー処理プログラム 214 の中で、コピー処理がどこまで進んだかを管理するために使用する表であり、複数のエントリからなる。各エントリは LU 131 を構成する実領域 132 に対応する。本実施形態では、IO 一時停止中に構成情報差分 222 を生成する。そのため構成情報 221 は 1 面しか持たない。その他の構成は図 12 と同様である。

【0095】

図 20 は、図 19 の仮想化スイッチ 11 における構成情報 221 の例を示す。

図 15 に比較して、LU のアドレスとオフセットの組は、Read コマンド用に 44 と 45、Write コマンド用に 46 と 47、及び 48 と 49 の 2 組からなる、計 3 組を持つ。本実施形態ではコピー処理プログラム 214 の進捗により、仮想ボリューム 100 の対応が Read 及び Write で異なる。さらに、Write の場合、LU 1311 と LU 1312 へ二重書きをする場合もあるため、Write 用のアドレスとオフセットの組は 2 組用意する。

【0096】

図 21 は、図 19 の仮想化スイッチ 11 における構成変更制御プログラム 212 の処理

を示す。

図16のフローチャートに比べて、ステップ627と603の間に、ステップ631～633および635が追加され、ステップ606と607の間にステップ634が追加されている。また、ステップ603'、612'の代わりにステップ603、612の処理が行われる。

【0097】

コピー処理631において、プロセッサ161は、コピー処理プログラム214の起動を行う。コピーが異常無しに行われた場合(632“Y”)、プロセッサ210はコピー進捗表224を参照して、構成情報差分222を作成する。尚、差分の作成についての詳細は図26を参照して後で説明する。

【0098】

ステップ634において、プロセッサ161はLU1311からLU1212へのコピーが全て完了したか否かを、コピー進捗表224を検査して判断する。コピーが完了している場合(634“Y”)、ステップ607を実行して処理を完了する。一方、コピーが完了していない場合(634“N”)には、プロセッサ161はステップ601から処理を繰り返す。

【0099】

コピーが異常終了した場合(632“N”)、プロセッサ161は回復処理を行う(635)。具体的には、構成変更制御プログラム212を実行する以前の構成情報221に戻すように差分222を作成し、全ての仮想化スイッチ11に構成変更要求を発行して完了を待つ。さらにプロセッサ161はステップ614以降の処理を実行し、管理コンソールへエラーを報告して処理を完了する。

【0100】

図22は、図21のフローチャートに示すコピー処理631の詳細な処理動作を示す。まずプロセッサ161は、構成変更制御プログラム212を起動してから最初のコピー処理6311が起動された場合(6311“Y”)、プロセッサ210に対しCOPY開始要求を送信し、コピー処理プログラム214を起動する(6312)。COPY開始要求に伴い、今回のコピー処理プログラム214の起動でコピーを行うアドレス範囲もプロセッサ210に指示する。

【0101】

一方、最初の起動でない場合(6311“N”)、プロセッサ210に対しCOPY再起動要求を送信し、コピー処理プログラム214を起動する(6313)。この場合も今回のコピー処理プログラム214の起動でコピーを行うアドレス範囲をプロセッサ210に指示する。いずれの場合も、プロセッサ161は次にイベント待ち状態となる(6314)。ここで、プロセッサ161が待つイベントは、コピー処理プログラム214の完了、後述するステップ718による、コピー優先要求またはコピー中断要求の受信である。コピー処理プログラム214の処理が完了した場合(6315“N”、6316“N”)、プロセッサ161はステップ632へ移る。

【0102】

コピー中断要求の場合(6315“Y”)、プロセッサ210に対しCOPY中断要求を送信し、コピー処理プログラム214を起動する(6317)。コピー優先要求の場合は、一旦コピー中断としてステップ6315“Y”から処理を行い、次にコピー処理を再起動する必要があるためプロセッサ161はステップ6316“Y”を実行して、ステップ6313以降の処理を繰り返す。その際、優先してコピーするアドレスがプロセッサ210から渡されるため、そのアドレスを含む範囲をステップ6313で指定する。ステップ6312および6313の通り、コピー処理プログラム214の起動は複数回に分けて行われる。さらにステップ6317ではコピーを一度中断させている。これにより入出力を一時保留する時間を短縮できる。

【0103】

図23は、図19の仮想化スイッチ11における構成管理プログラム211の処理を示

す。

本実施形態では、I/O一時保留状態の間に構成情報221の面切替えを行わないため、構成管理プログラム211の処理は第一の実施形態に示したものに近い。そこで、図23を図7のフローチャートと比較するに、図7と相違する点は、ステップ718を追加したところにある。

【0104】

本実施形態では、構成管理プログラム211が呼び出される契機は、図7の場合に加え、I/O処理プログラム213からのコピー優先要求またはコピー中断要求による。いずれの場合も、プロセッサ210はステップ709Nと判断することになり、さらにプロセッサ161に対し、コピー優先要求またはコピー中断要求を行う(718)。ただし、コピー優先要求の場合、優先してコピーするアドレスがI/O処理プログラム213から渡されるので、そのアドレスもプロセッサ161に渡す。

このように、ステップ718の処理により構成管理プログラム211はI/O処理プログラム213の処理による要求を構成変更制御プログラム212の処理へ反映する働きをする。

【0105】

図24は、図19の仮想化スイッチ11におけるI/O処理プログラム213の処理を示す。

図18に示すフローチャートに比べて、ステップ805'の後にステップ806が追加され、ステップ820の後にステップ821が追加されている。

【0106】

ステップ806において、プロセッサ210は、構成管理プログラム211を呼び出し、受け付けた入出力がWriteの場合はコピー優先要求を行う。その際、Writeの場合は受け付けた入出力でWriteされるアドレス範囲を構成管理プログラム211に渡す。

【0107】

本実施形態において、タイマ215のタイムアウト(820“Y”)はコピー処理によるので、ステップ821において、プロセッサ210は構成管理プログラム211を呼び出してコピー中断要求を行う。

【0108】

ステップ806および821の処理によって、一時保留した入出力によるコピー処理を中断または優先させることで、その入出力が保留されている時間を短縮できる。なお、ステップ806において、受け付けた入出力がReadの場合はコピー中断要求を行うことにすると、さらに入出力を一時保留する時間を短縮することができる。(この場合、ステップ821は不要になる)。但し、その分コピー処理は遅くなる。

【0109】

図25は、図19の仮想化スイッチ11におけるコピー処理プログラム214の処理を示す。

コピー処理プログラム214が起動する契機は、構成変更制御プログラム212からのCOPY開始要求およびCOPY再起動要求、COPY中断要求と、データを実際にコピーする入出力の完了のタイミングである。コピー開始要求があった場合(901“Y”)、プロセッサ210はコピー進捗表224の全エントリを、未コピー状態を示す値(この例では「0」)に初期化する(902)。そして構成変更制御プログラム212から渡されたアドレス範囲に対応するコピー進捗表のエントリを、コピー中を示す値(この例では「1」)にセットする(903)。

【0110】

さらに、プロセッサ210はコピー進捗表224で「1」(コピー中)にセットした領域に対するコピーを行う入出力命令(以下、COPY I/O)を発行する(904)。具体的には、コピーを行う入出力要求を生成し、出力キュー251へエンキューする。さらにプロセッサ210はイベント待ちを行って(905)、コピー処理プログラム214の

起動契機のうち、COPY開始要求以外のものを待ち、ステップ906以降の処理を実行する。また、コピー再起動要求があった場合(906“Y”)、プロセッサ210はステップ903以降の処理を繰り返す。

【0111】

COPY IOが完了した時(907“Y”)、プロセッサ210はこの入出力の結果が正常終了であることを確認し(908“Y”)、コピー進捗表224を更新する(909)。具体的には、COPY IOが終了した領域に対するエントリにコピー済みを示す値(この例では「4」)をセットし、コピー中断を示す値(この例では「2」)が格納されているエントリに「0」(未コピー)をセットする。さらに、コピー進捗表224のエントリで、「1」(コピー中)がセットされているエントリが無い場合(910“Y”)、プロセッサ210は構成変更制御プログラム212成功を返して処理を完了する(911)。コピー中のエントリがある場合(910“N”)、プロセッサ210はステップ904から処理を繰り返す。

【0112】

COPY中断要求を受けた場合(912“Y”)、プロセッサ210は、コピー進捗表224の「1」(コピー中)がセットされている全エントリへ「2」(コピー中断)をセットし(913)、ステップ905から繰り返す。

COPY IOが異常終了した場合には(908“N”)、プロセッサ210は、コピー進捗表224の「1」(コピー中)および「2」(コピー中断)がセットされている全エントリに「0」(未コピー)をセットする(914)。そして、構成変更制御プログラム212にエラーとその要因を返して、終了する(915)。

このように、コピー処理プログラム214ではステップ904のCOPY IOの処理によってデータのコピーを行うことができる。

【0113】

図26を参照して、第四の実施形態によるデータ移行の動作原理を説明する。

図26(A)は、コピー処理631中における仮想ボリューム100と実領域132との対応関係、およびコピー進捗表224との対応関係を示す。

【0114】

実領域132から仮想ボリューム100への矢印は、仮想ボリューム100に対するRead要求での対応関係を表しており、仮想ボリューム100から実領域132への矢印は、仮想ボリューム100に対するWrite要求での対応関係を表している。点線になっている矢印は、これに対する入出力が保留されていることを示している。

【0115】

なお、仮想ボリューム上の領域1001から実領域13211と実領域13221へ矢印が出ているが、これは1001にWrite要求があったとき、2個の実領域13211、13221に同じデータが二重に書き込まれることを意味している。

さらに、実領域13212、13213から実領域13222、13223への矢印はこの方向でコピーが行われていることを示している。コピー進捗表の2241a、2242a、2243a、2244aは、仮想ボリューム上の領域1001、1002、1003、1004に対応している。

【0116】

図26(A)では、2241aはコピー済み「4」になっており、領域100-1について実領域13211から実領域13221へのコピーが完了していることを示している。コピー済みの領域に関しては、Readは、移行先即ち1311から、Writeは移行元および移行先の両方、即ち1311と1312の両方へ書き込まれる。これは、LU1312への移行が途中で失敗した場合でも、最新のデータがLU1311に残るようにするためである。

【0117】

また、2242a~2244aはコピー中「1」になっており、現在これらに対応する実領域13212~13214から実領域13222~13224へのデータのコピーが

行われていることを示している。2245aは未コピー「0」となっており、実領域13215から実領域13225へのコピーは未着手であることを示している。

【0118】

図26(B)は、領域1004に対し、ホストプロセッサ12がWrite要求を発行した後のコピー進捗表224を示している。コピー中の領域に対するWriteを受信すると、IO処理プログラム213のステップ806(図24)によりプロセッサ210はコピー優先要求を行う。すると、プロセッサ161はまずコピー処理631のコピー中断処理6317を行うため、その結果プロセッサ210がコピー処理プログラム214のステップ912を実行することにより、コピー進捗表224のコピー中「1」の領域にコピー中断の印「2」をセットするため、表224bの状態になる。

【0119】

次にプロセッサ210はCOPY IOの完了を待って、ステップ909を実行し、コピー進捗表224を更新する。ここで、実領域13222へのコピーのみが完了しているとすると、これに対応するエントリ2242bはコピー中断「2」からコピー済み「4」へ、さらにその他のエントリはコピー中断の印「2」から未コピーの印「0」へ更新されることになる。さらにプロセッサ161はコピー処理631におけるステップ6313の処理を行い、領域1004に対する実領域13214のアドレスを指定してプロセッサ210がコピー処理プログラム213のステップ906以降の処理を行わせるため、領域1004に対応する2244bにコピー中の印「1」をセットする。このようにして図26(C)の状態になる。つまり、プロセッサ210は実領域13214から13224へのコピーを優先して行う。

尚、その他の処理は、第三の実施形態と同様なので説明を省略する。

【0120】

以上に述べたように、システムの稼動中に仮想ボリューム100の割当て先を、一方のLU1311から他方のLU1312に移行することが可能になる。

【0121】

次に、図27から図28を用いて、第五の実施形態によるを説明する。

図27はストレージシステムの全体構成を示す図である。

図1に示したシステムに比べて、独立した構成変更制御装置16が無くなり、代わって各仮想化スイッチ11内に構成変更制御プログラム16'を備える。また、ストレージ装置135はコピー制御部136を備える。

【0122】

ストレージ装置135はストレージ装置13と異なり、通信部260を備えてLAN15と接続しており、さらにコピー制御部136を備える。これにより自装置135内のLU131aから別のLU131cへデータをコピーすることができる。ストレージ装置135にあるコピー制御部136は、コピー元LU131へのホストプロセッサ12からの入出力を許可したままコピー処理を行えること(以下、オンライン中にコピー可能という)が望ましい。勿論オンライン中にコピー可能でなくても良い。このオンライン中にコピー可能でない場合、入出力を一時保留している時間が長くなることになる。

【0123】

なお、どのLU131からどのLU131へコピー処理を行うか、例えばLU131aからLU131bへコピーする、の指定およびコピー処理の開始、完了の要求は、コピー制御部136が通信部260を経由して、構成変更制御装置16または管理コンソール14から受け取る。

【0124】

図28は、第五の実施形態における構成変更制御プログラム212の処理を示すフローチャートである。

図16に比べ、ステップ627の後に、ステップ640の処理が挿入されたことが相違する。

【0125】

ステップ640の処理で、プロセッサ161は、コピー制御部136に対し、LU131aからLU131bへのコピー処理の完了を依頼し、その完了報告を待つ。オンライン中にコピー可能な場合は、ステップ640の前なら任意の時点で、システム管理者がコピー開始の依頼を行うことができる。なお、オンライン中にコピー可能でない場合、ステップ640でプロセッサ161はコピー制御部136に対し、コピー処理の開始の依頼を、完了の依頼に先立ち行う。コピー開始の依頼を早めておくことで、ステップ640で待つ完了報告が早く返ってくることが期待できる。

【0126】

尚、プロセッサ161及び210は第三の実施形態と同様の処理を行い、仮想ボリューム100を、LU131aからLU131bに対応するように構成を変更する。その他の処理は第三の実施形態と同様なので説明を省略する。

【0127】

第五の実施形態によれば、ストレージ装置135が持つコピー機能を利用することにより、仮想化スイッチ11にLU131をコピーする機能がなくても、システムの稼動中に仮想ボリューム100の割当て先を、一方のLU131aから他方のLU131bに移行することができる。

【0128】

次に、図29を用いてストレージシステムの更に他の例（第六の実施形態）を説明する。
この例は、ストレージ装置137が仮想化機能を備えたものである。各ストレージ装置137は、図27に示すストレージ装置13、及びコピー制御部136、通信部260を有し、更に仮想化スイッチ11を備えている。各仮想化スイッチ11は、前述の図27と同様に構成変更制御プログラム16'を内蔵して、構成変更制御の機能を実現している。その他の構成は、前記第五の実施形態と同様であるので、説明を省略する。

【0129】

この例によれば、冗長構成の仮想化スイッチによるシステムの稼動中に、仮想ボリューム100の割当て先を、一方のLU131aから他方のLU131bに移行することが可能なストレージ装置を実現できる。

【0130】

以上、いくつかの実施形態について説明したが、本発明は上記の例に限定されずに、種々変形して実施し得る。例えば、図12及び図19の例において、プロセッサ161または210のいずれか一方を省略し、残る他方のプロセッサで、プログラムの処理等を兼用することができる。

【図面の簡単な説明】

【0131】

【図1】 第一の実施形態によるストレージシステム全体の構成を示す図。

【図2】 図1の仮想化スイッチ11の内部構成を示す図。

【図3】 図1の構成変更制御装置16の内部構成を示す図。

【図4】 ストレージシステムにおける構成情報221のテーブルを示す図。

【図5】 図1のストレージシステムにおける構成変更制御装置16と仮想化スイッチ11との通信手順を示す図。

【図6】 構成変更制御装置16における構成変更制御プログラム212の処理を示すフローチャート。

【図7】 仮想化スイッチ11における構成管理プログラム211の処理を示すフローチャート。

【図8】 仮想化スイッチ11におけるIO処理プログラム213の処理を示すフローチャート。

【図9】 第二の実施形態による仮想化スイッチ11の内部構成を示す図。

【図10】 第二の実施形態における構成変更制御プログラム212の処理を示すフローチャート。

- 【図 1 1】 図 1 0 における調停処理 6 0 0 の処理動作を示すフローチャート。
【図 1 2】 第三の実施形態による仮想化スイッチ 1 1 の内部構成を示す図。
【図 1 3】 第三の実施形態における構成変更制御装置 1 6 と仮想化スイッチ 1 1 との通信手順を示す図。
【図 1 4】 図 1 2 における一時保留制御表 2 2 3 の一例を示す図。
【図 1 5】 図 1 2 における構成情報 2 2 1 のテーブルを示す図。
【図 1 6】 図 1 2 における構成変更制御プログラム 2 1 2 の処理を示すフローチャート。
【図 1 7】 図 1 2 における構成管理プログラム 2 1 1 の処理を示すフローチャート。
【図 1 8】 図 1 2 における I O 処理プログラム 2 1 3 の処理を示すフローチャート。
【図 1 9】 第四の実施形態における仮想化スイッチ 1 1 の内部構成を示す図。
【図 2 0】 図 1 9 における構成情報 2 2 1 のテーブルを示す図。
【図 2 1】 図 1 9 における構成変更制御プログラム 2 1 2 の処理を示すフローチャート。
【図 2 2】 図 2 1 のフローチャートにおけるコピー処理 6 3 1 の詳細を示すフローチャート。
【図 2 3】 図 1 9 における構成管理プログラム 2 1 1 の処理を示すフローチャート。
【図 2 4】 図 1 9 における I O 処理プログラム 2 1 3 の処理を示すフローチャート。
【図 2 5】 図 1 9 におけるコピー処理プログラム 2 1 4 の処理を示すフローチャート。
【図 2 6】 第四の実施形態におけるデータ移行の動作原理を説明する図。
【図 2 7】 第五の実施形態によるストレージシステムの全体構成を示す図。
【図 2 8】 第五の実施形態における構成変更制御プログラム 2 1 2 の処理を示すフローチャート。
【図 2 9】 第六の実施形態によるストレージシステムの全体構成を示す図。

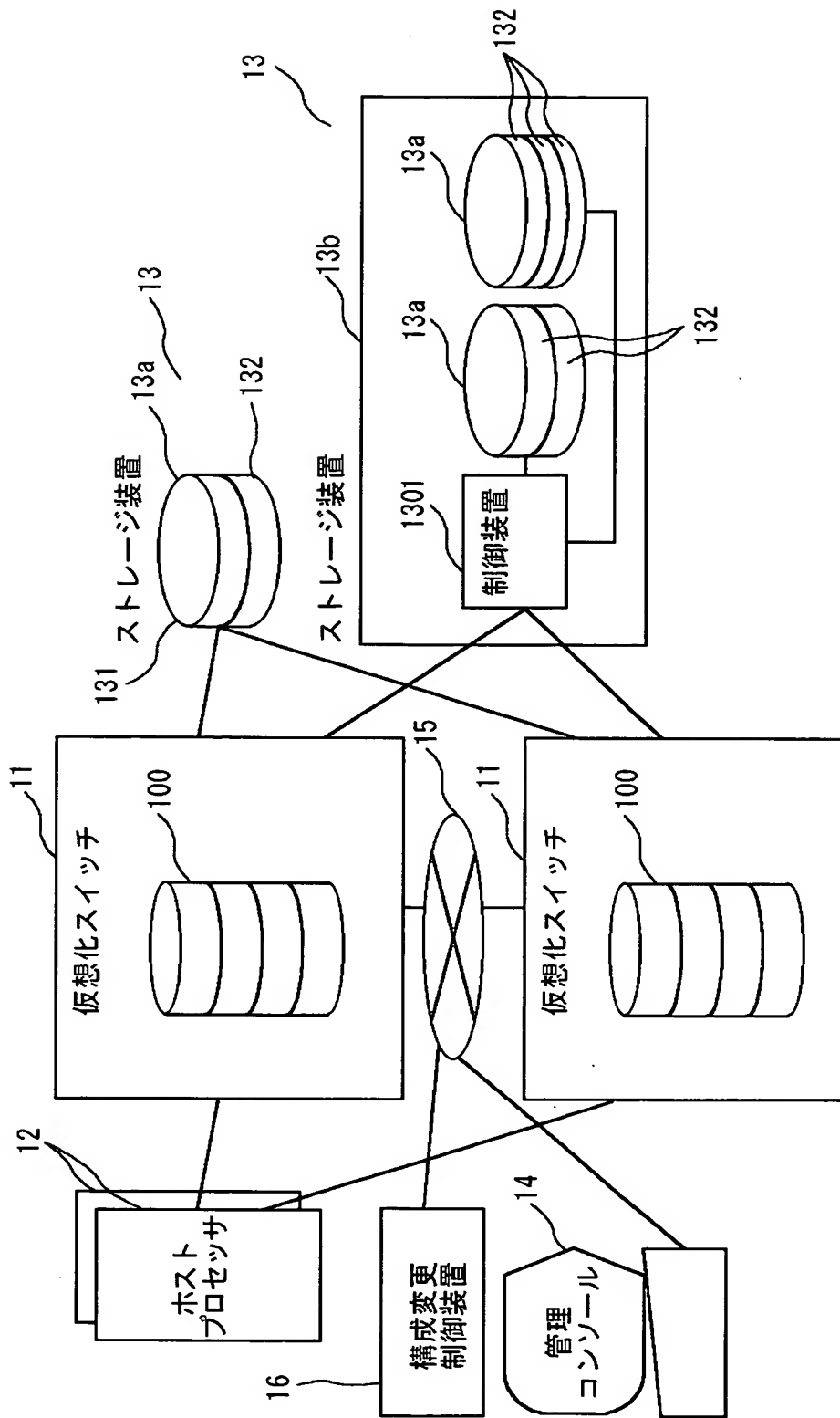
【符号の説明】

【0 1 3 2】

1 1	仮想化スイッチ、	1 2	ホストプロセッサ
1 3	ストレージ装置、	1 3 1	L U
1 4	管理コンソール、	1 6	構成変更制御装置
1 6 1、2 1 0	プロセッサ、	2 1 1	構成管理プログラム
2 1 2	構成変更制御プログラム、	2 1 3	I O 処理プログラム
2 1 4	コピー処理プログラム、	2 2 1	構成情報
2 4 2	保留中キュー、	2 5 2	処理中キュー

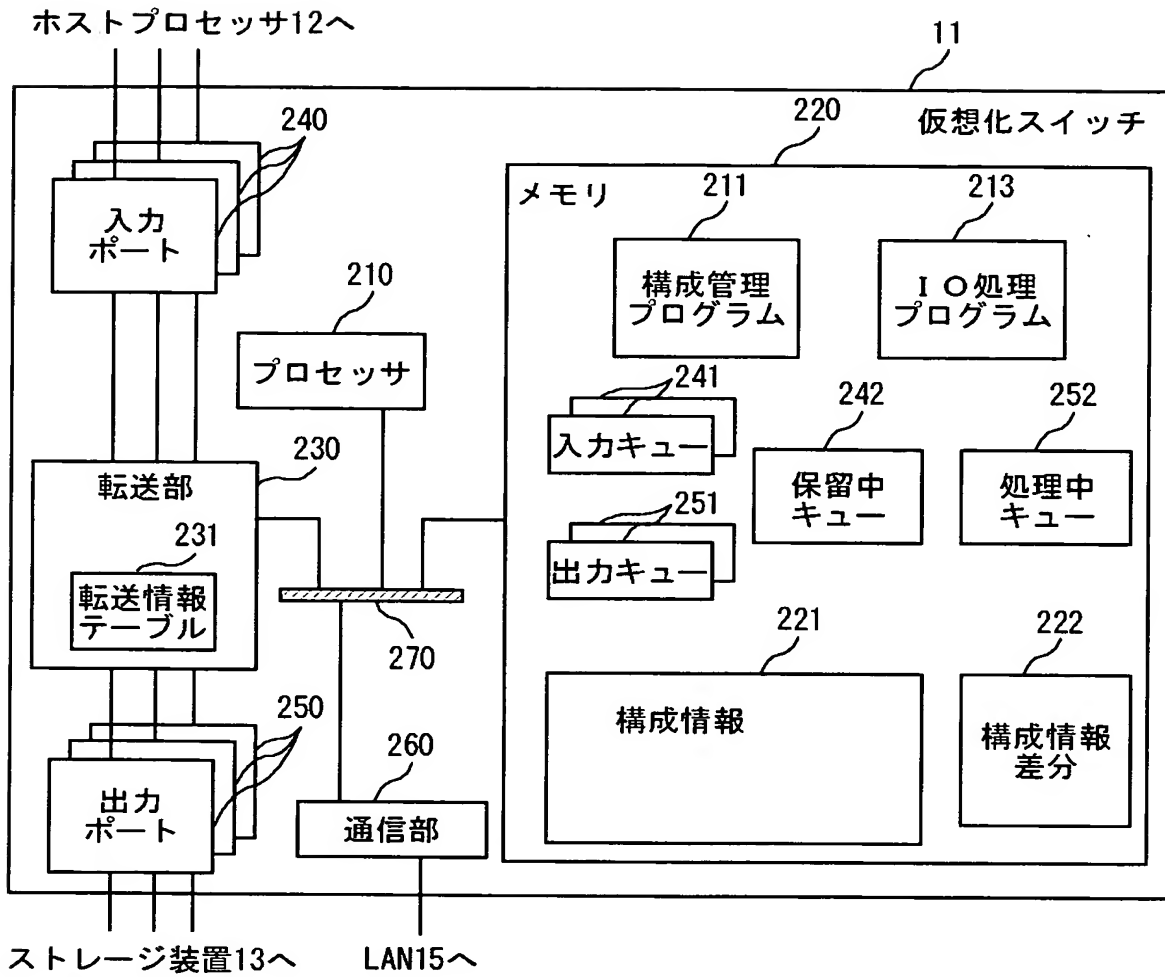
【書類名】 図面
【図 1】

図 1



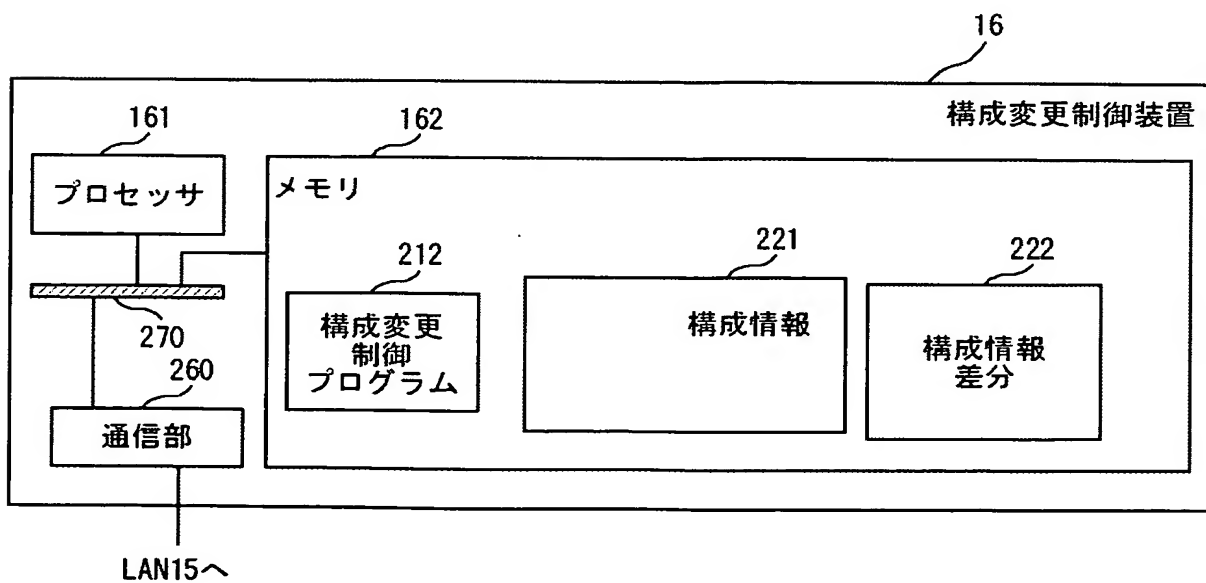
【図 2】

図 2



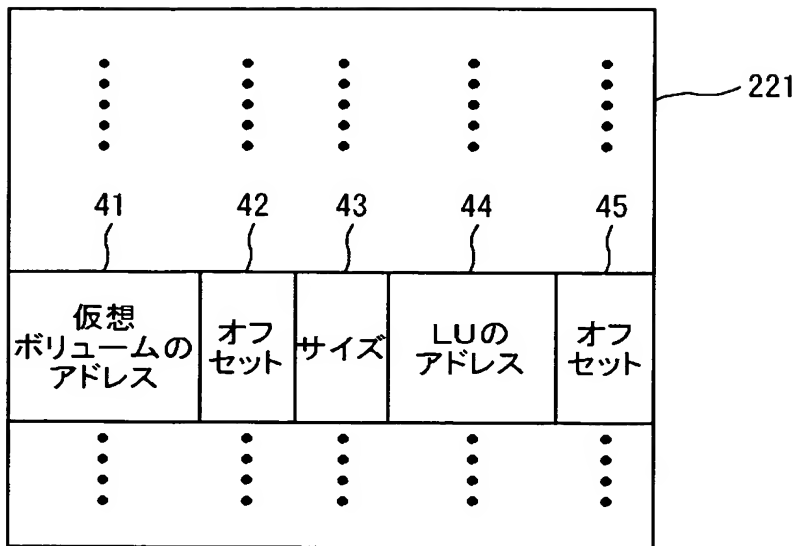
【図 3】

図 3



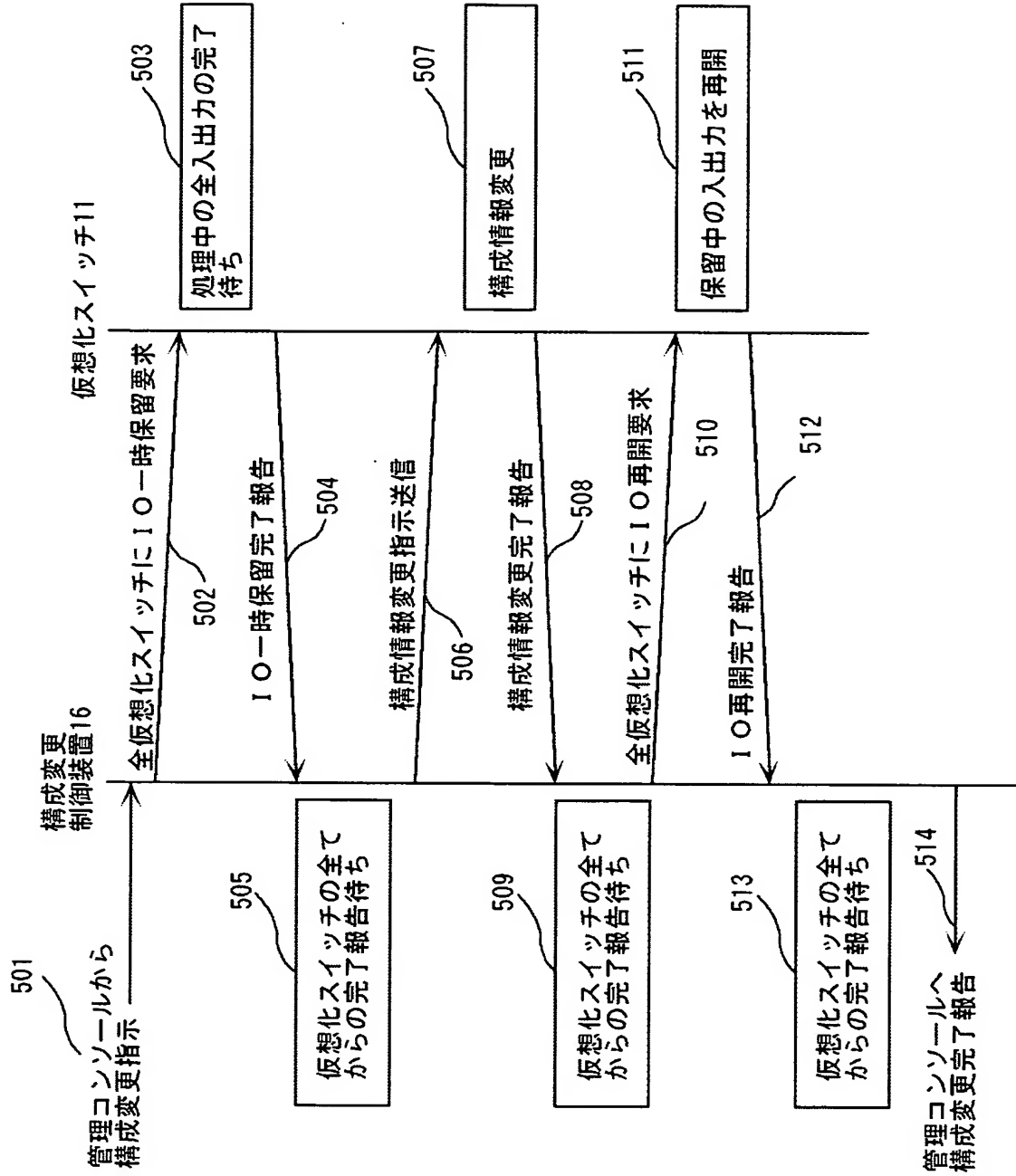
【図 4】

図 4



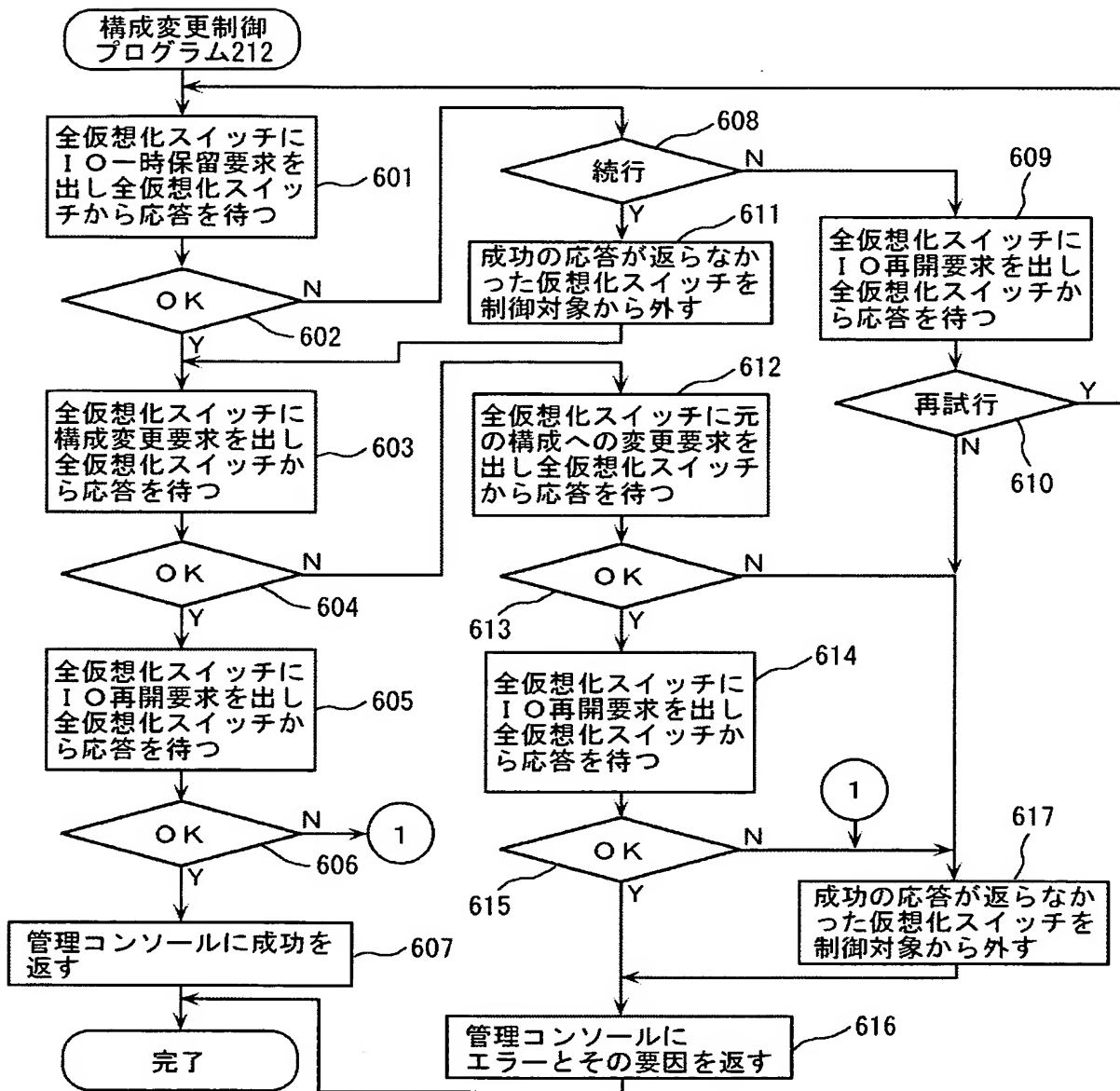
【図 5】

図 5



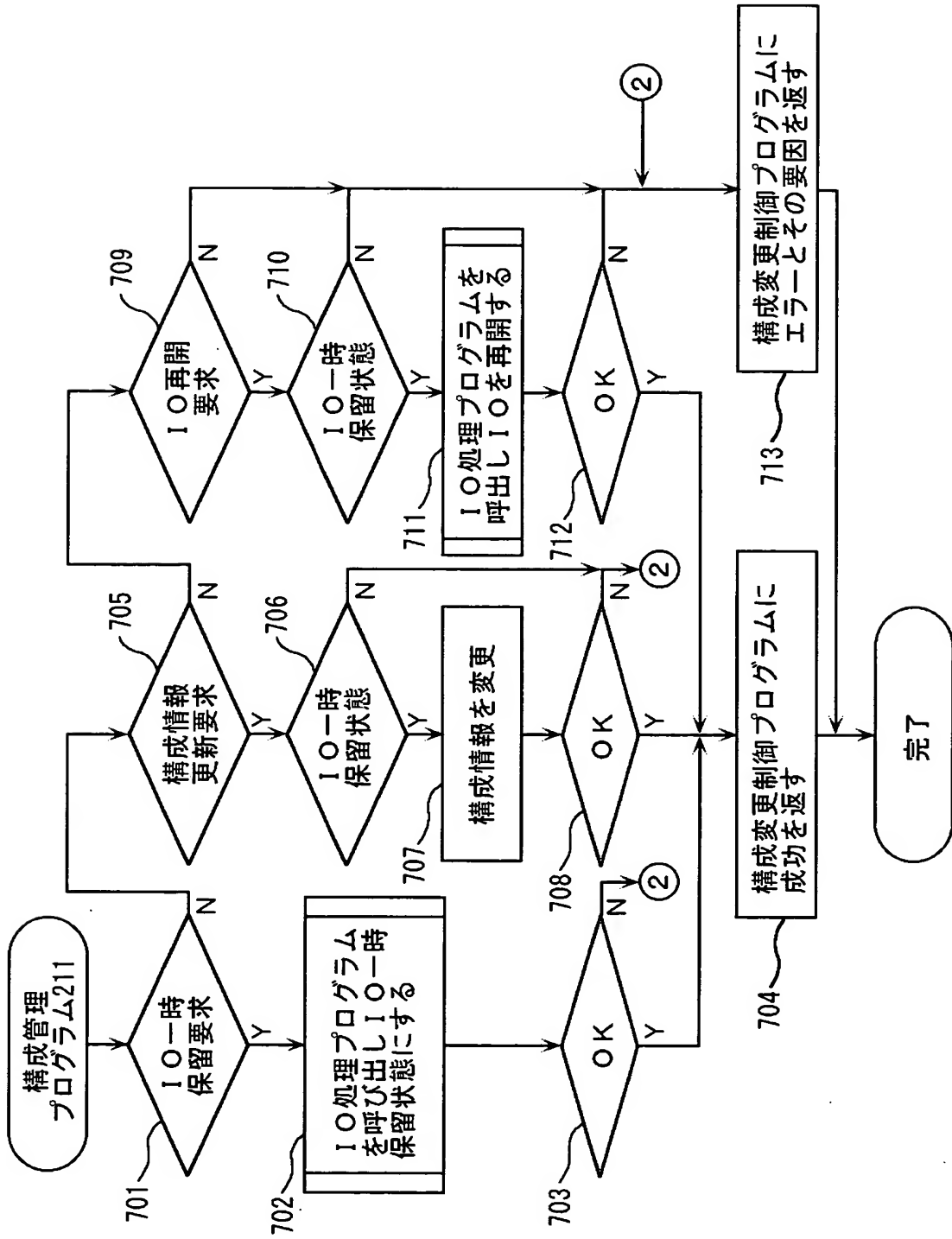
【図 6】

図 6



【図 7】

図 7



【図8】

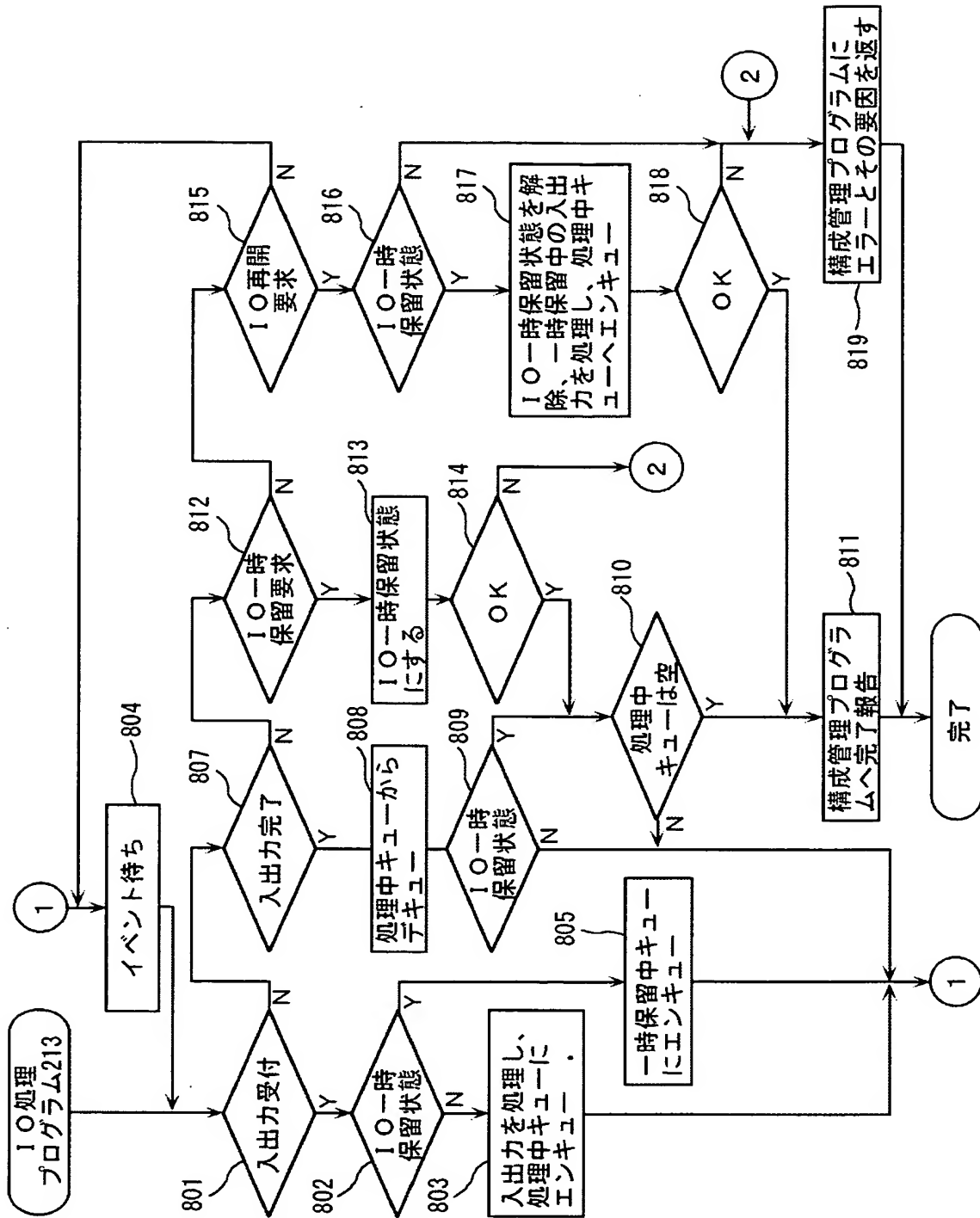
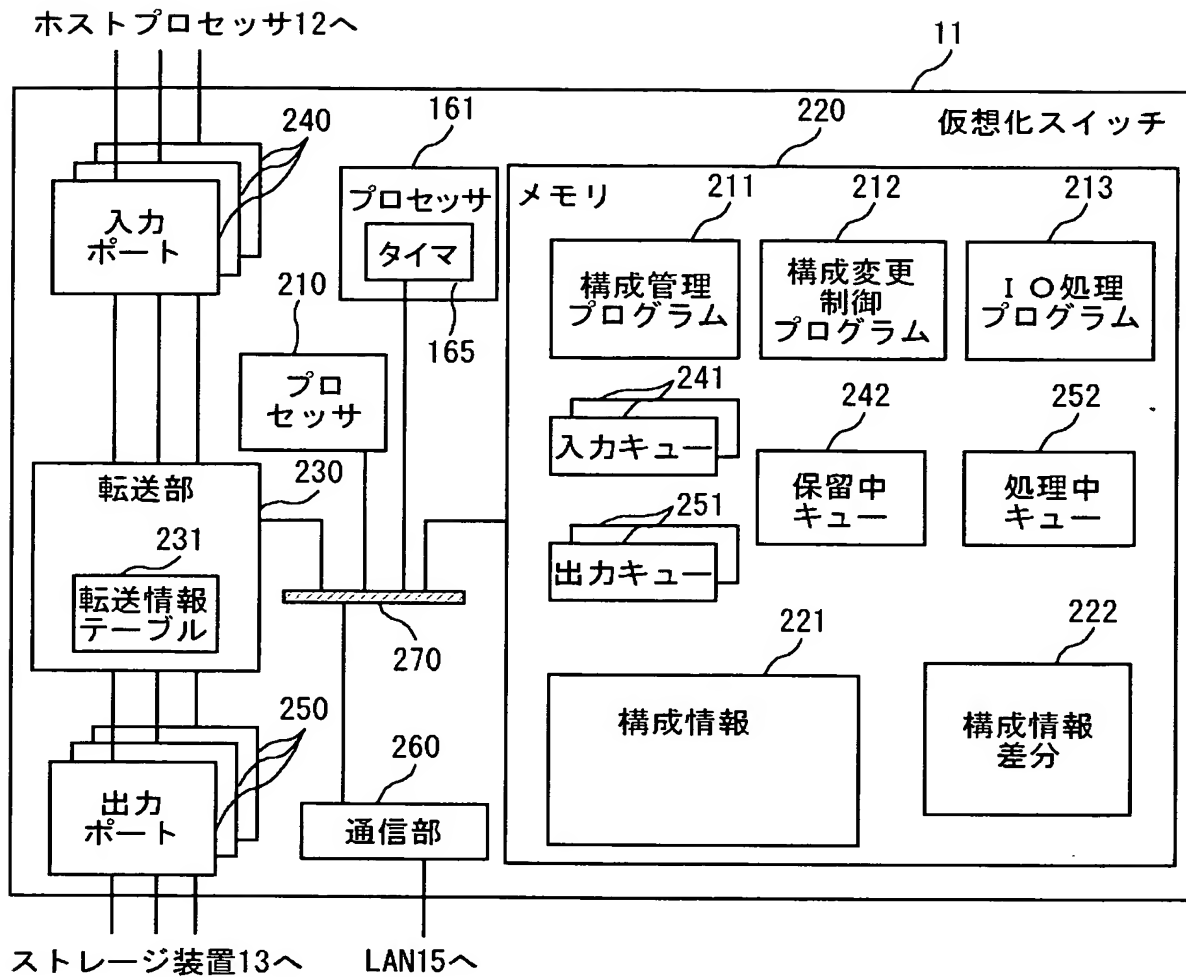


図 8

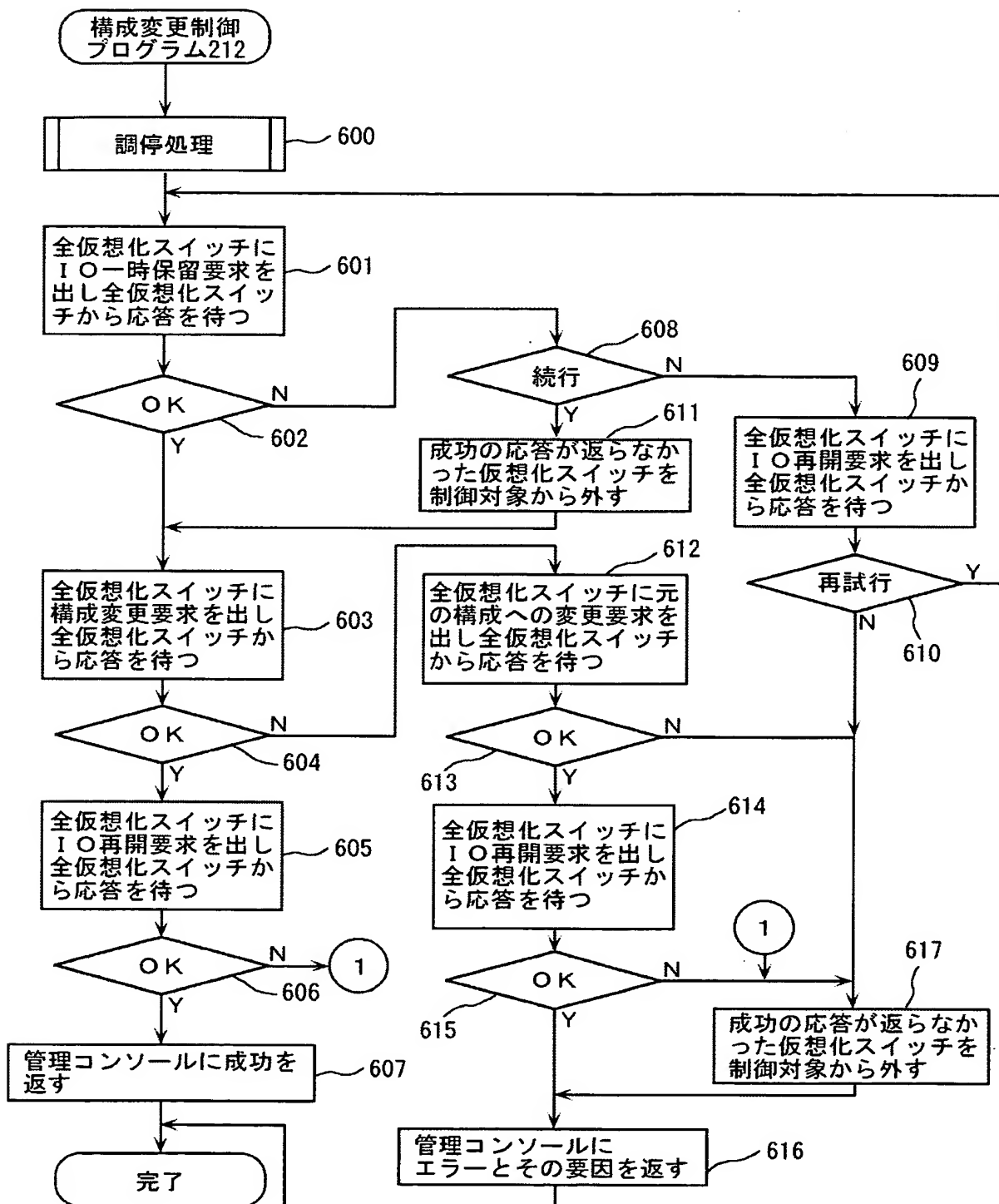
【図 9】

図 9



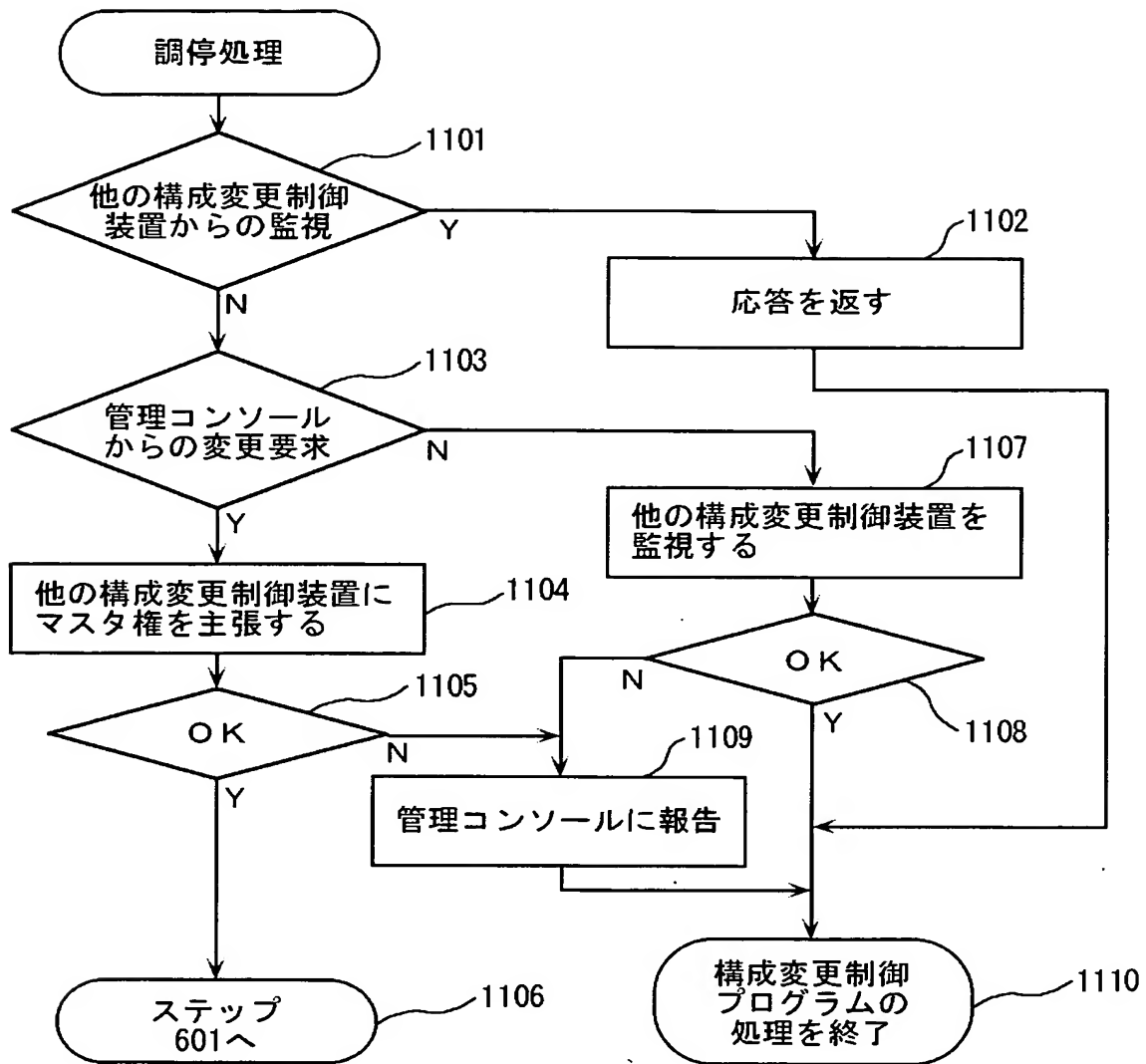
【図 10】

図 10



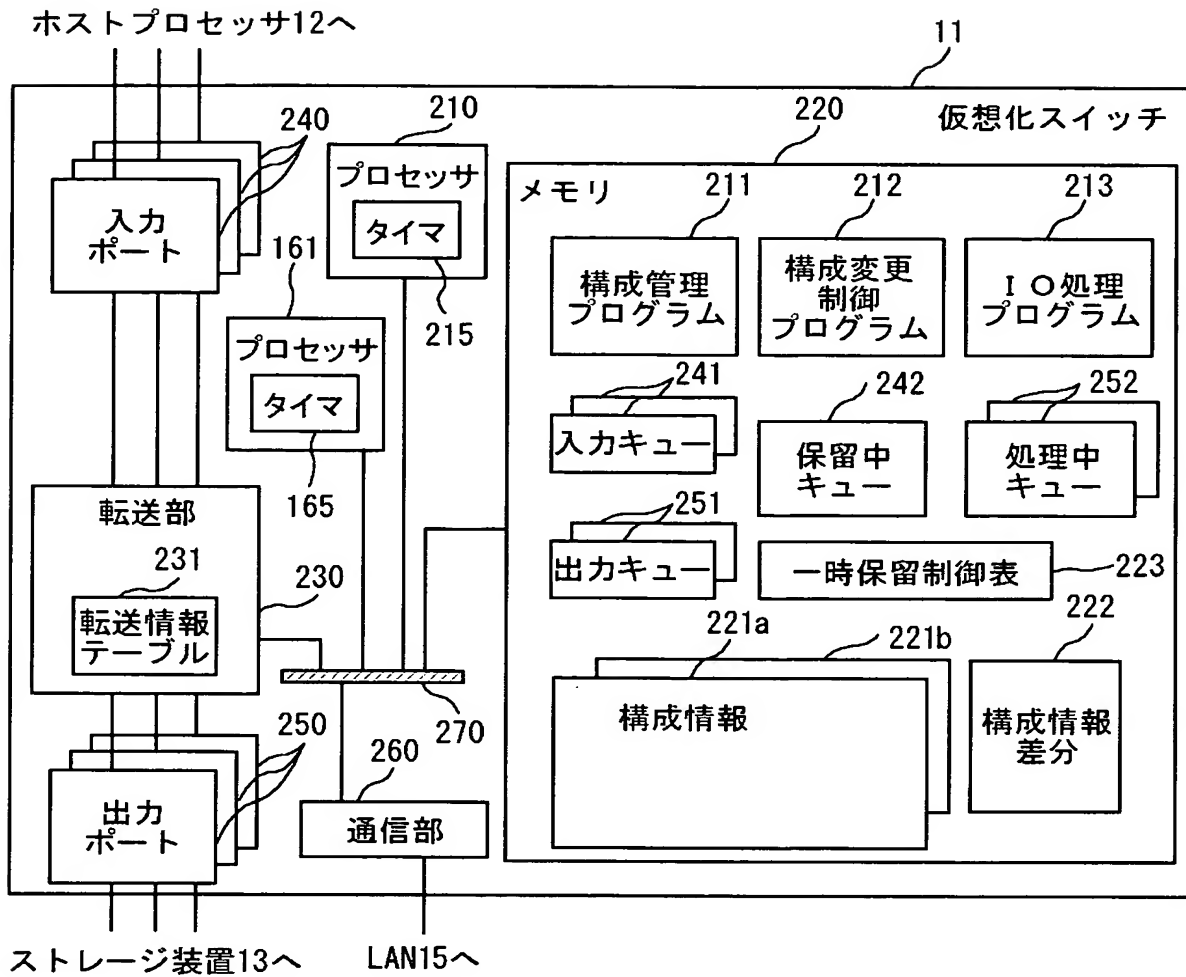
【図 11】

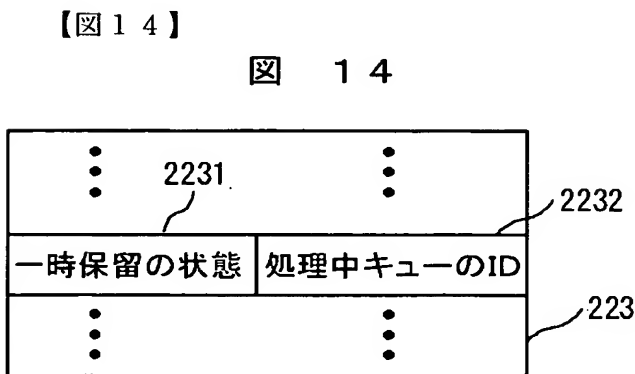
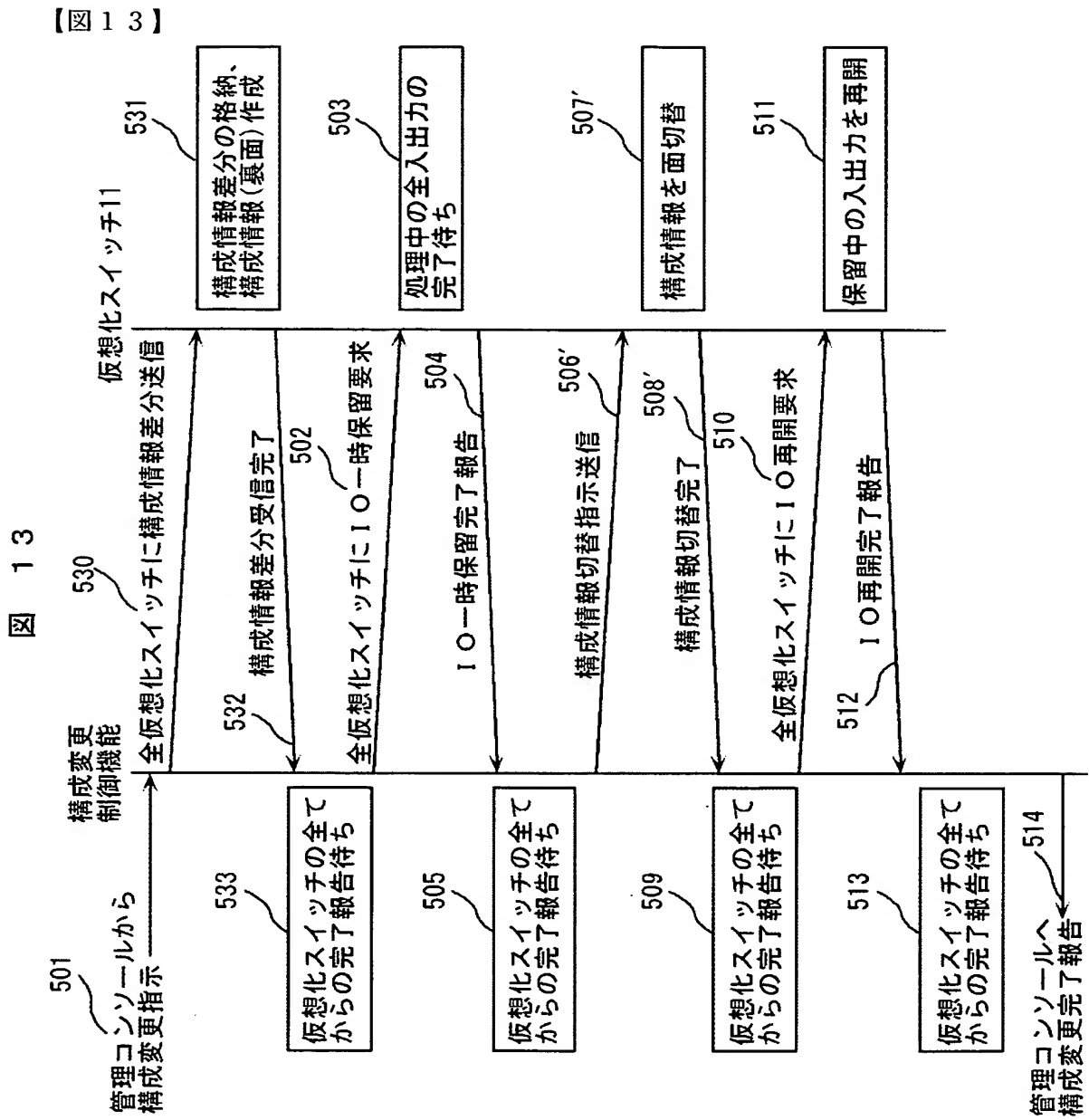
図 11



【図 12】

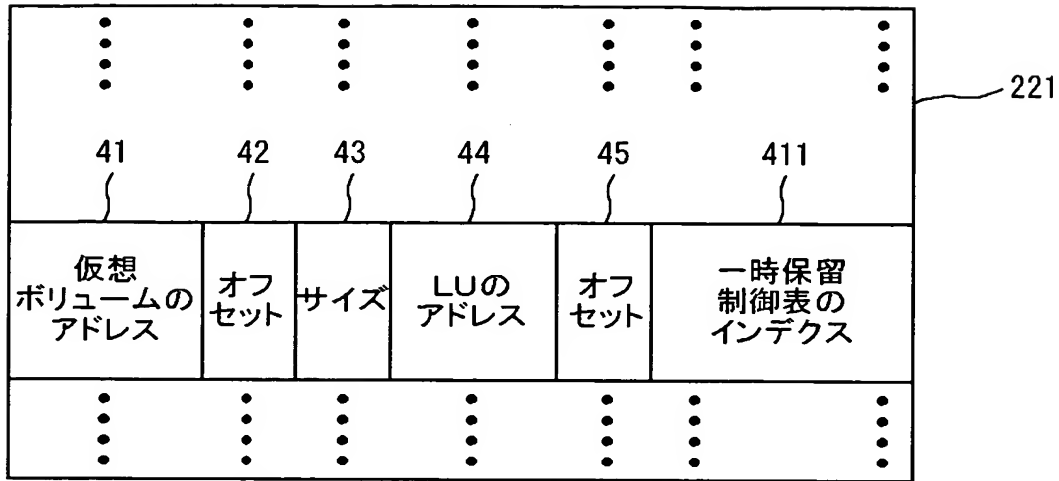
図 12





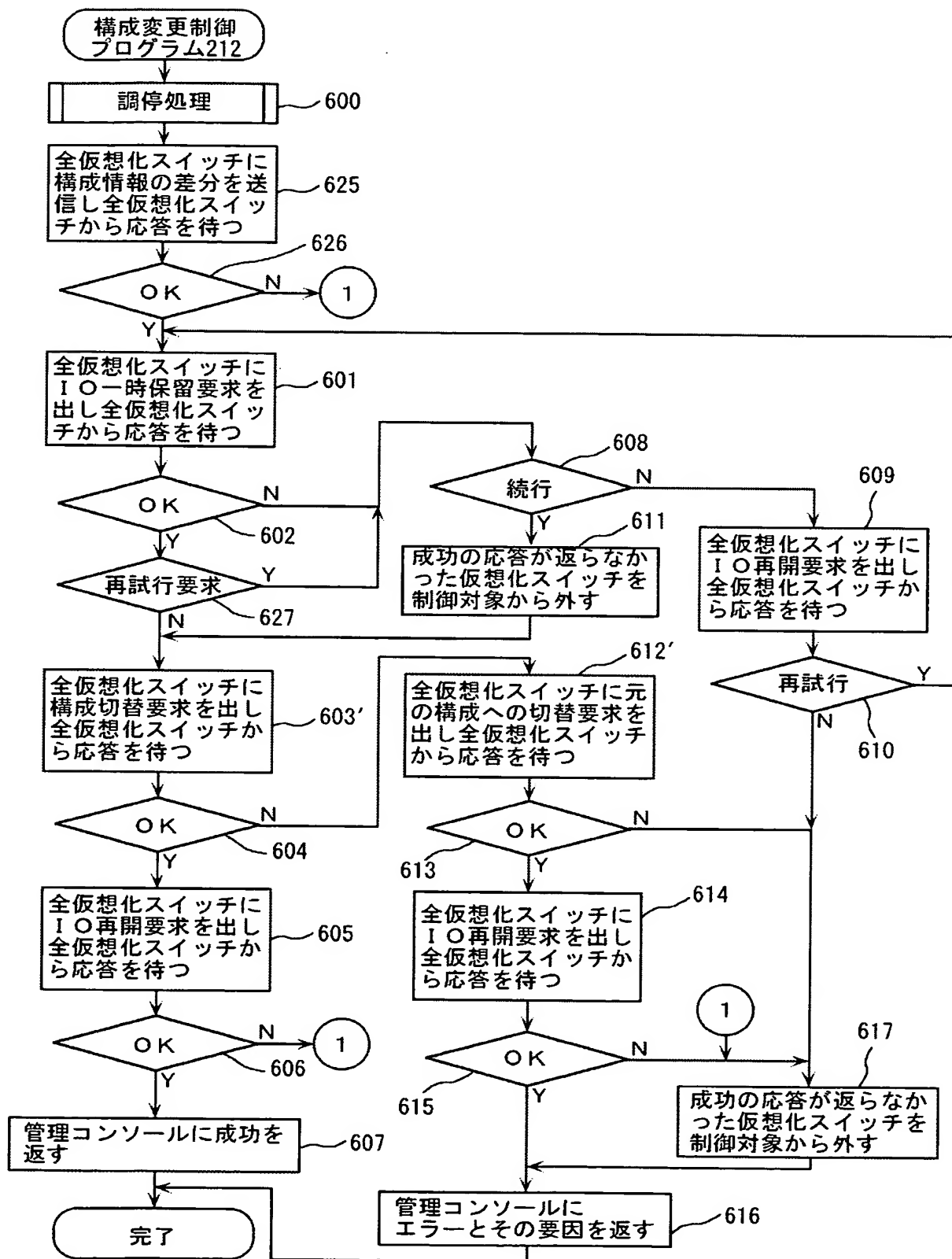
【図 15】

図 15



【図 16】

図 16



【図 17】

図 17

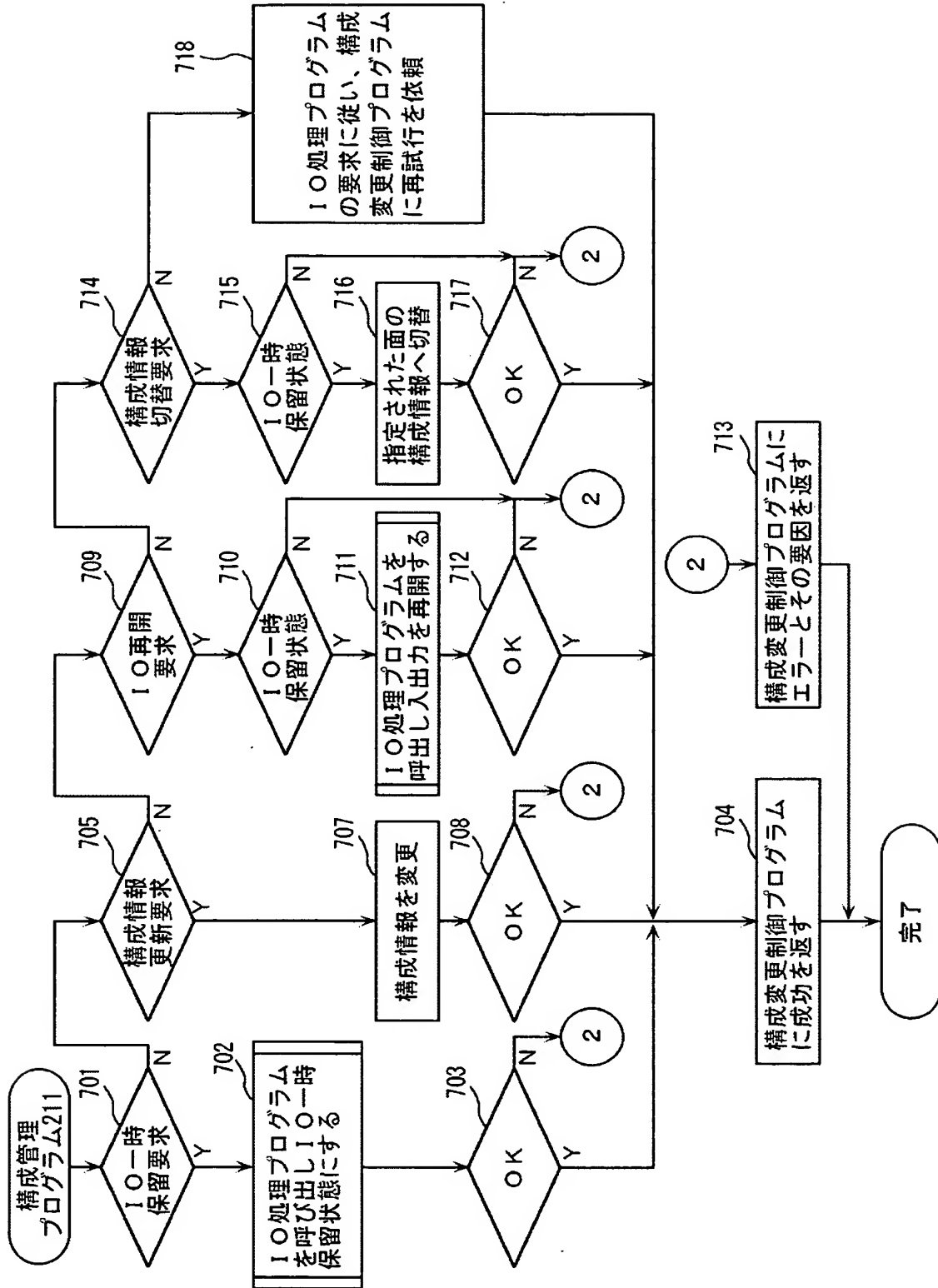
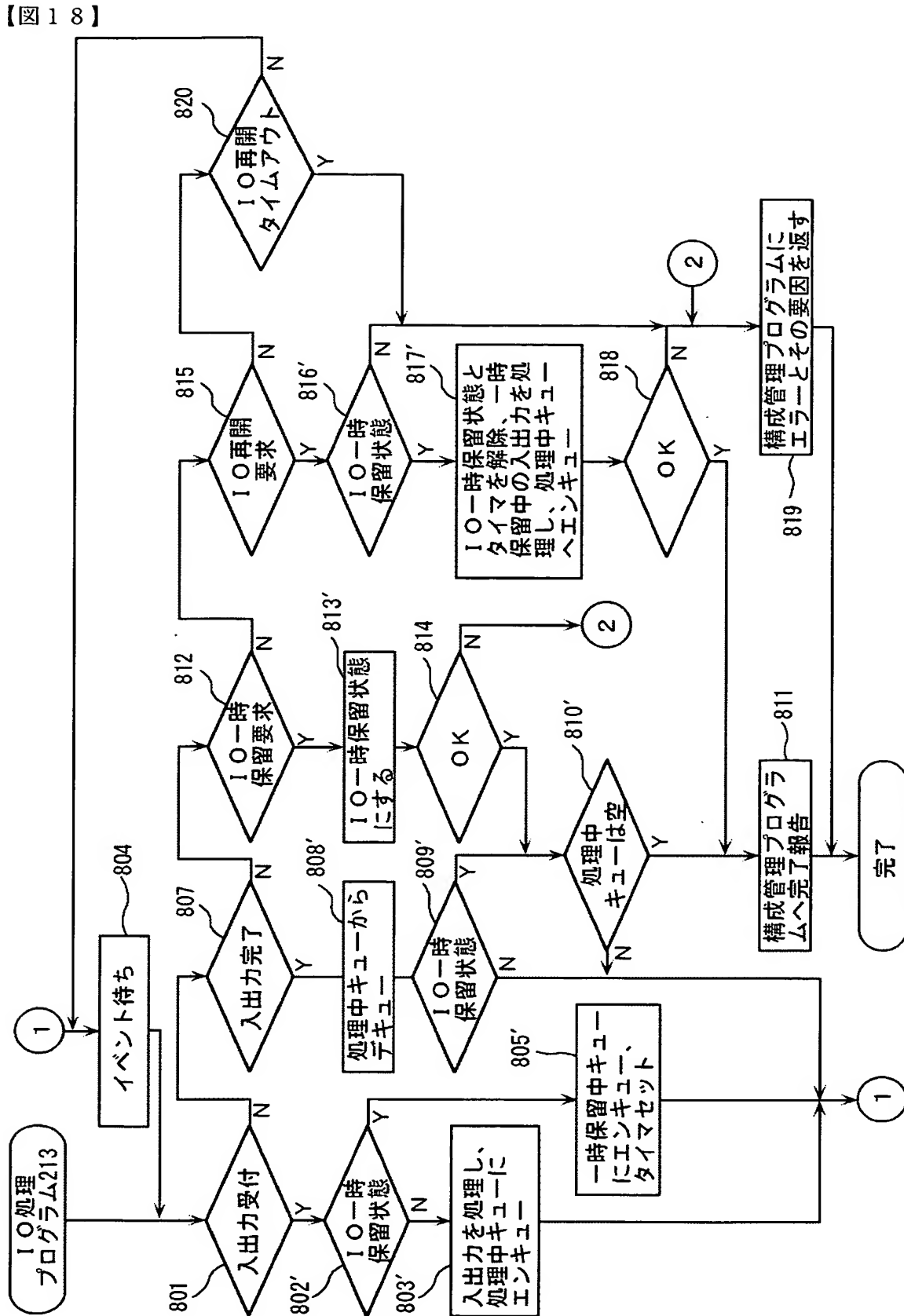
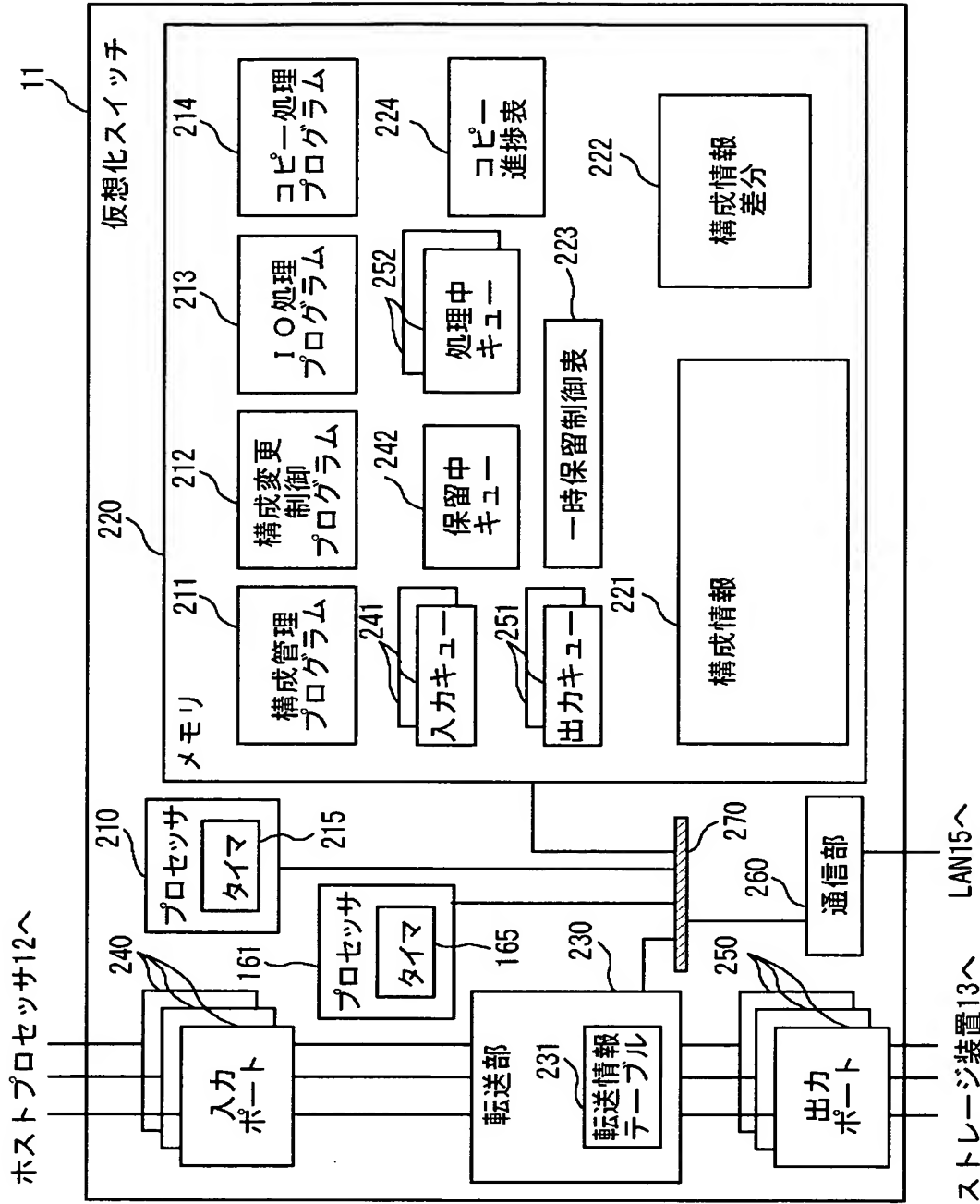


図 18

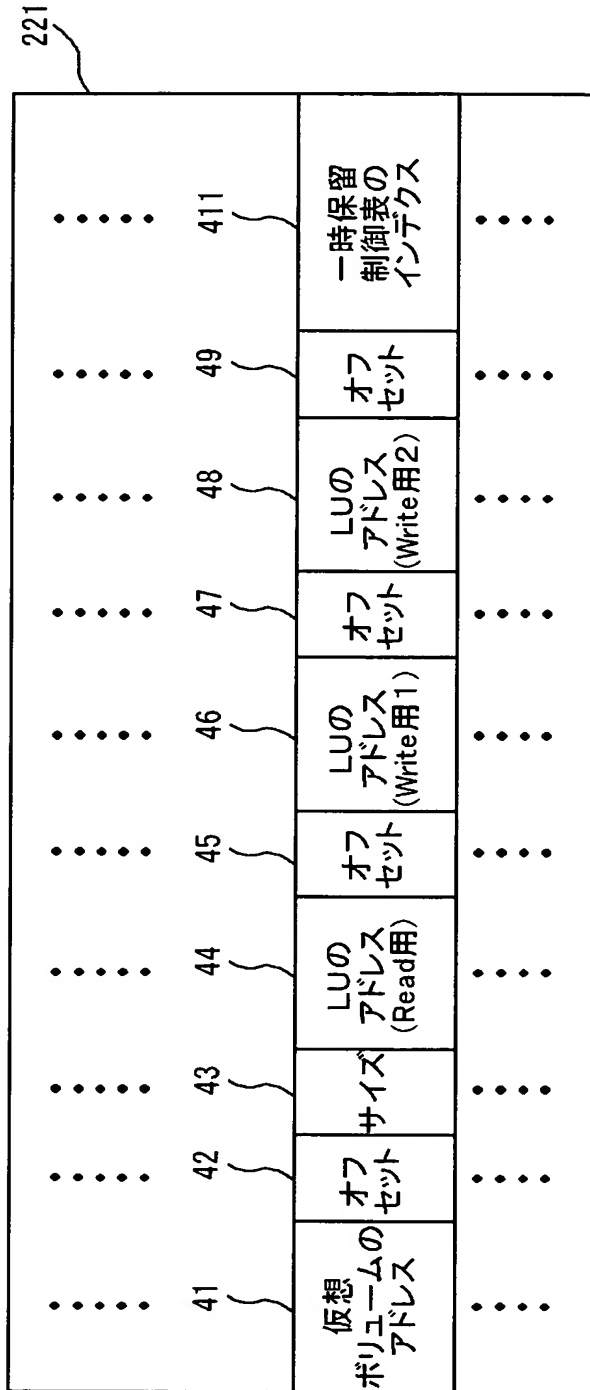


【図19】

図 19



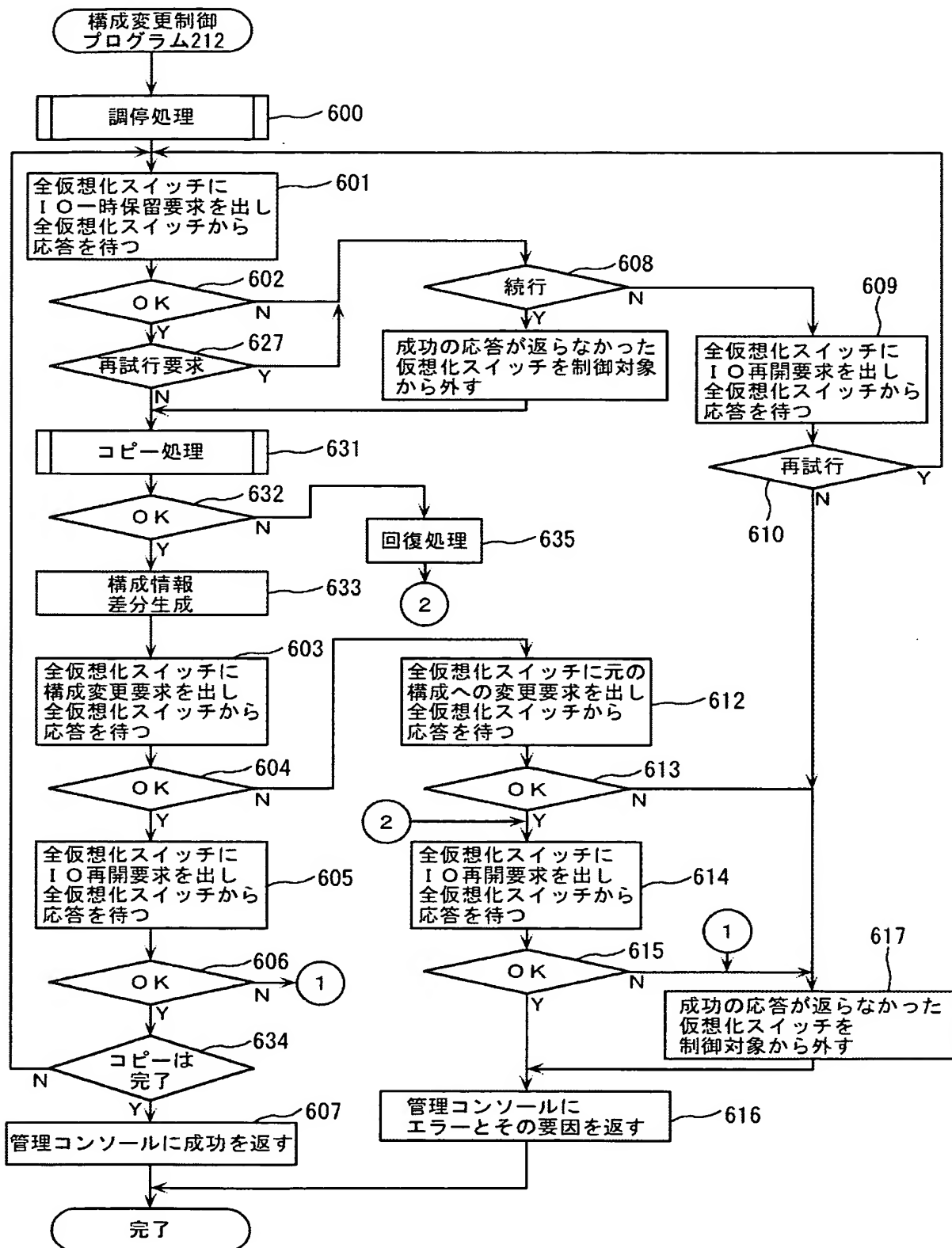
【図 20】



20

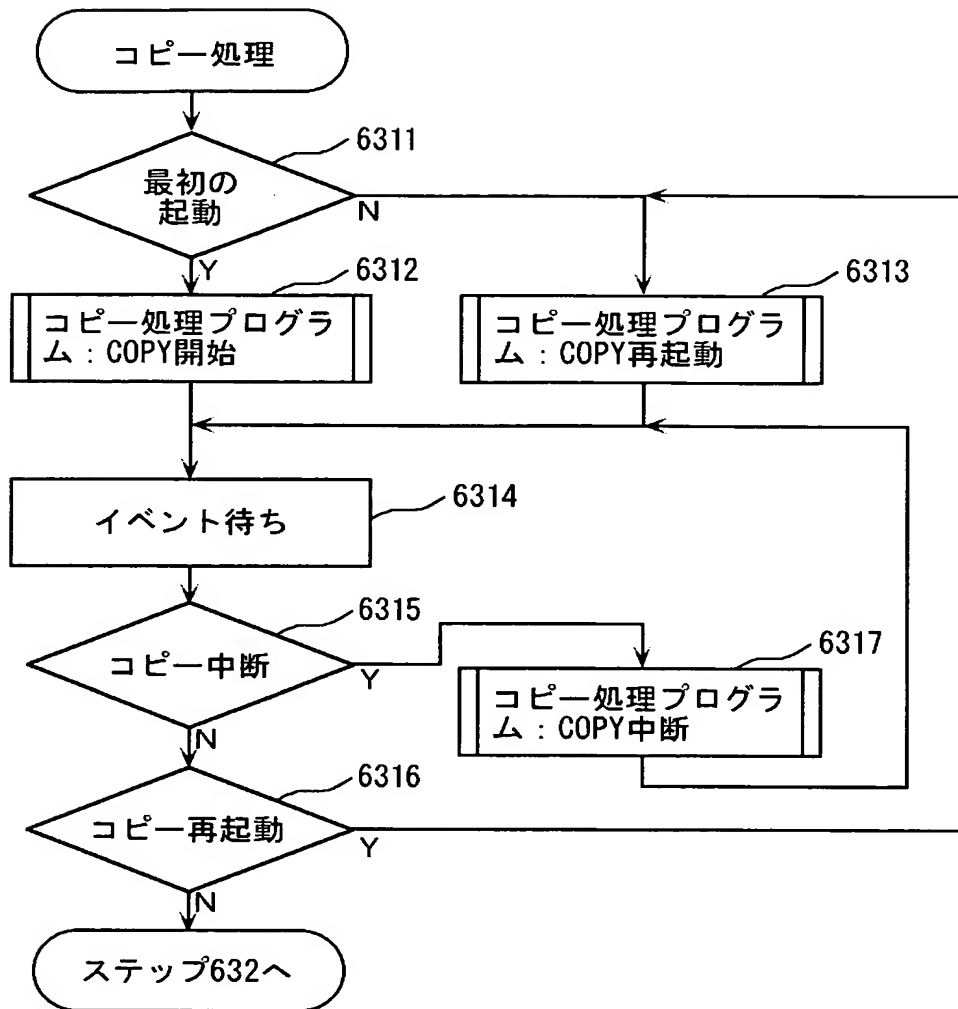
【図 21】

図 21



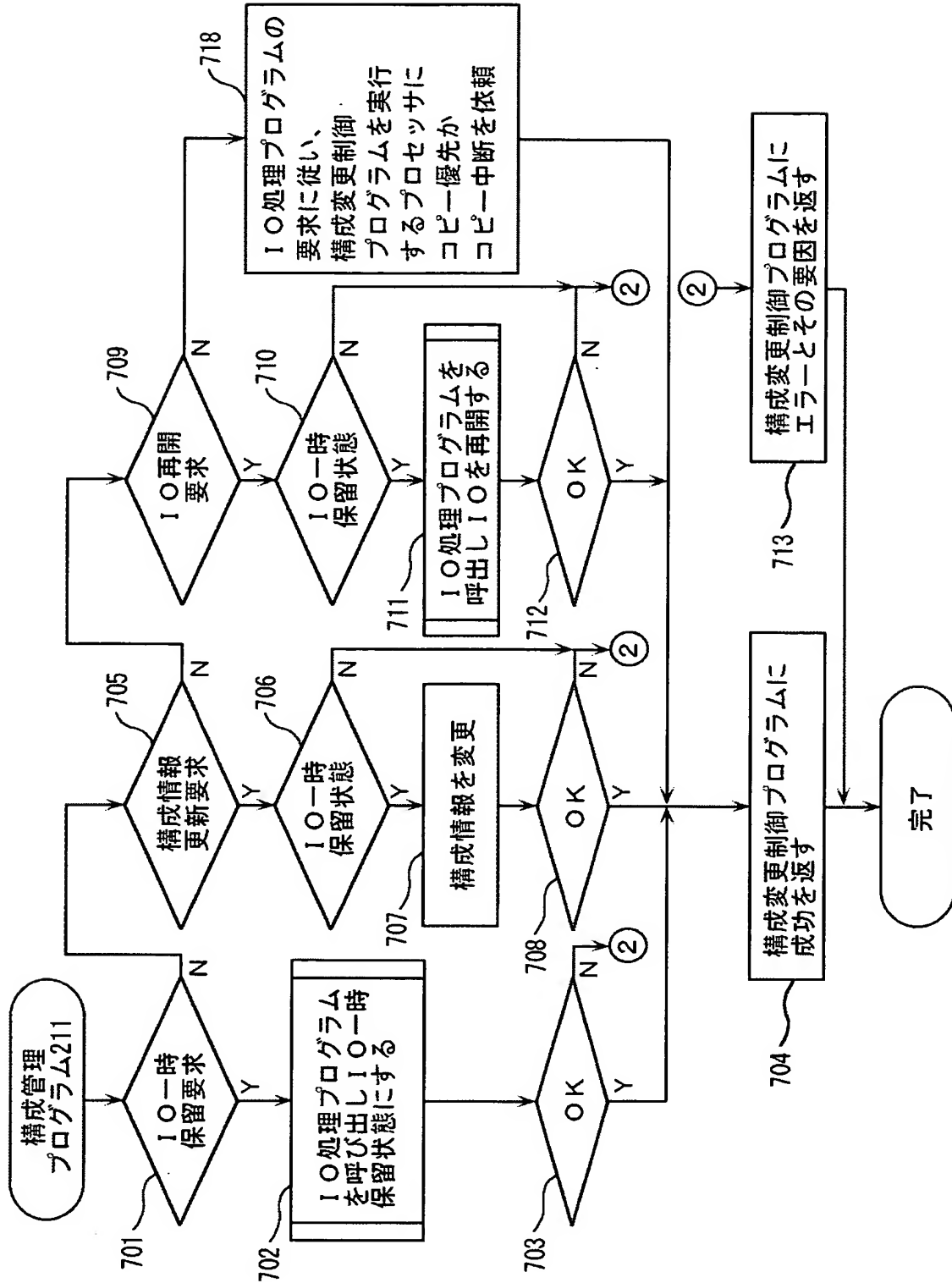
【図 22】

図 22

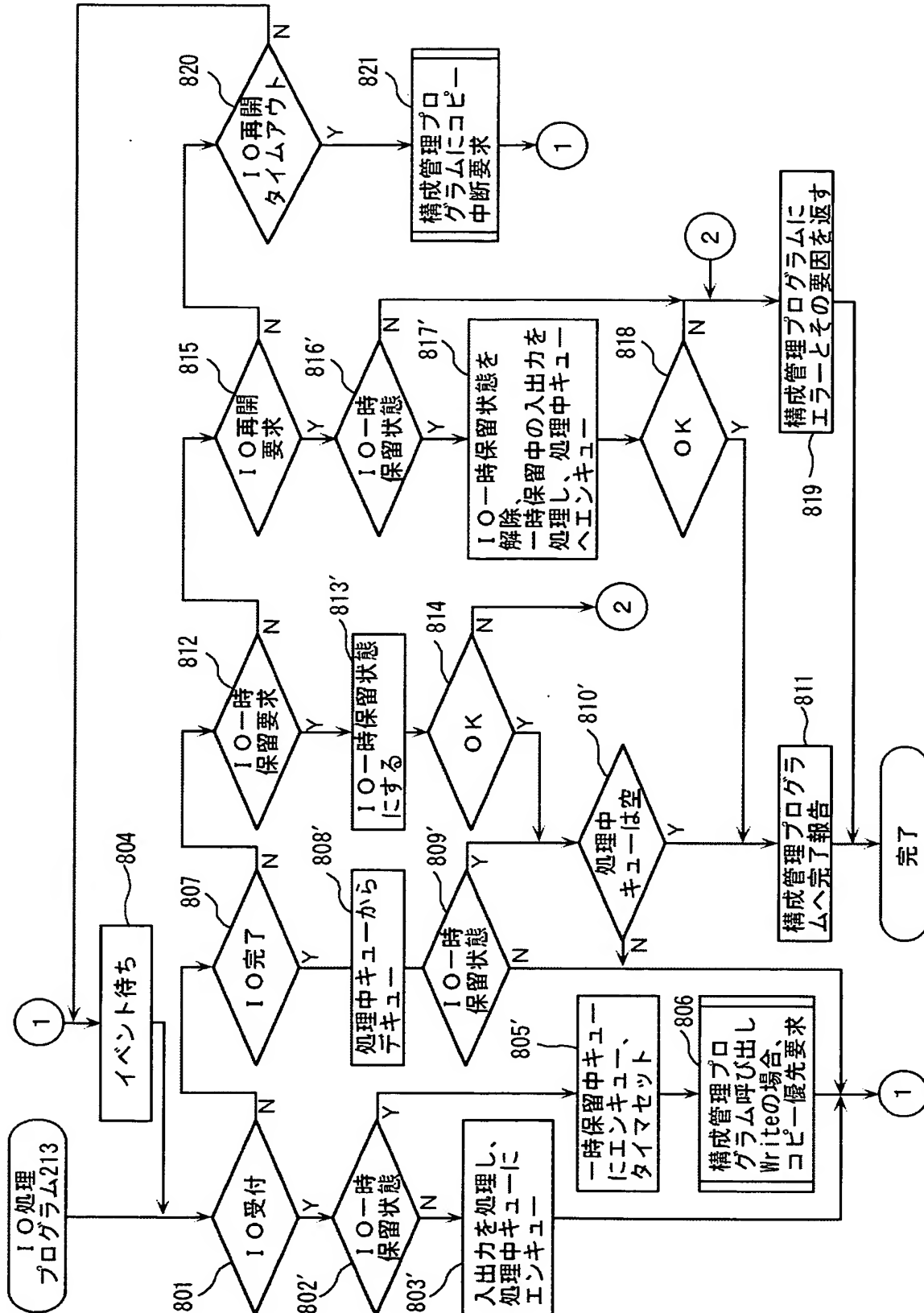


【図23】

図 23

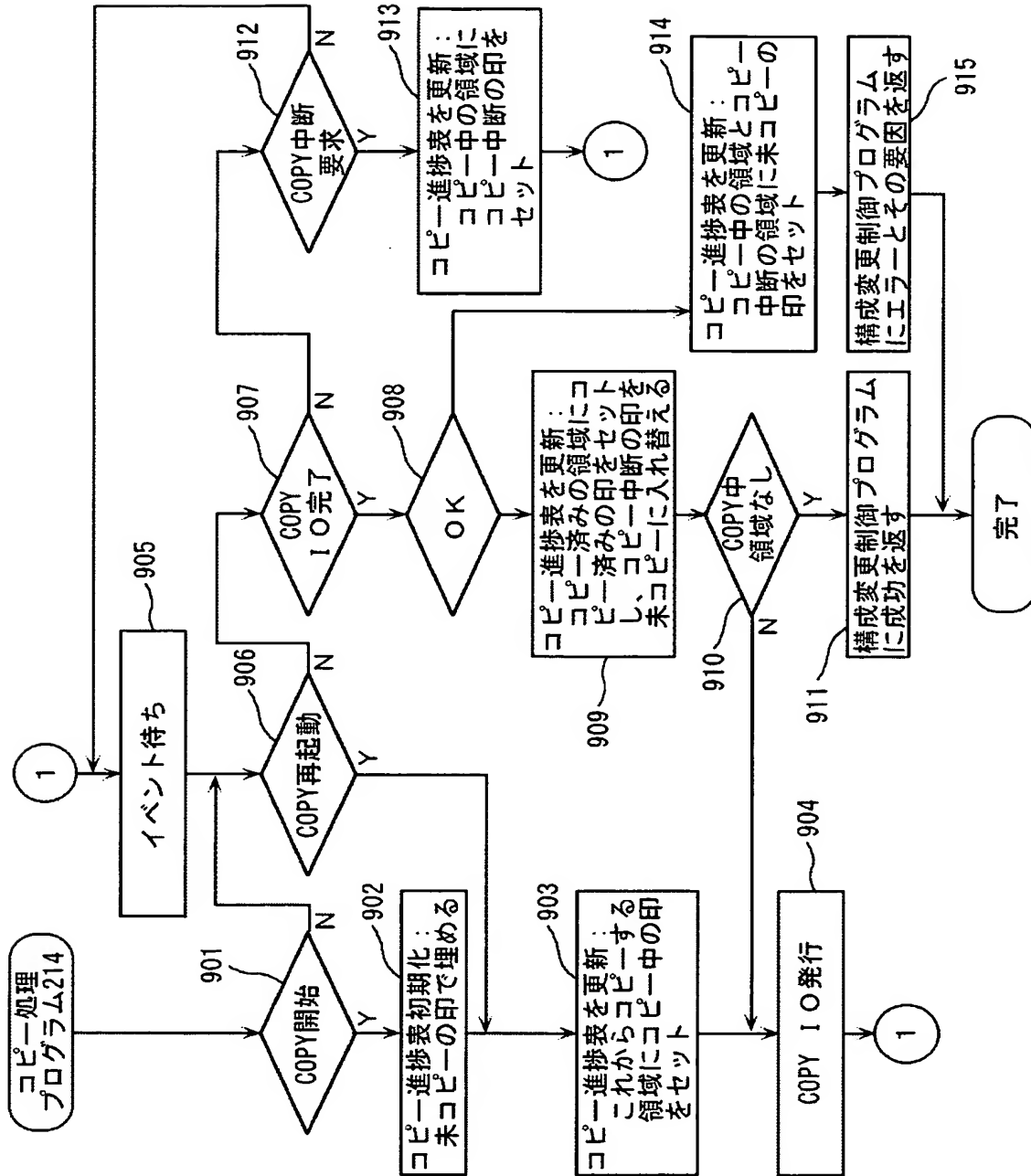


【図 2 4】



【図 25】

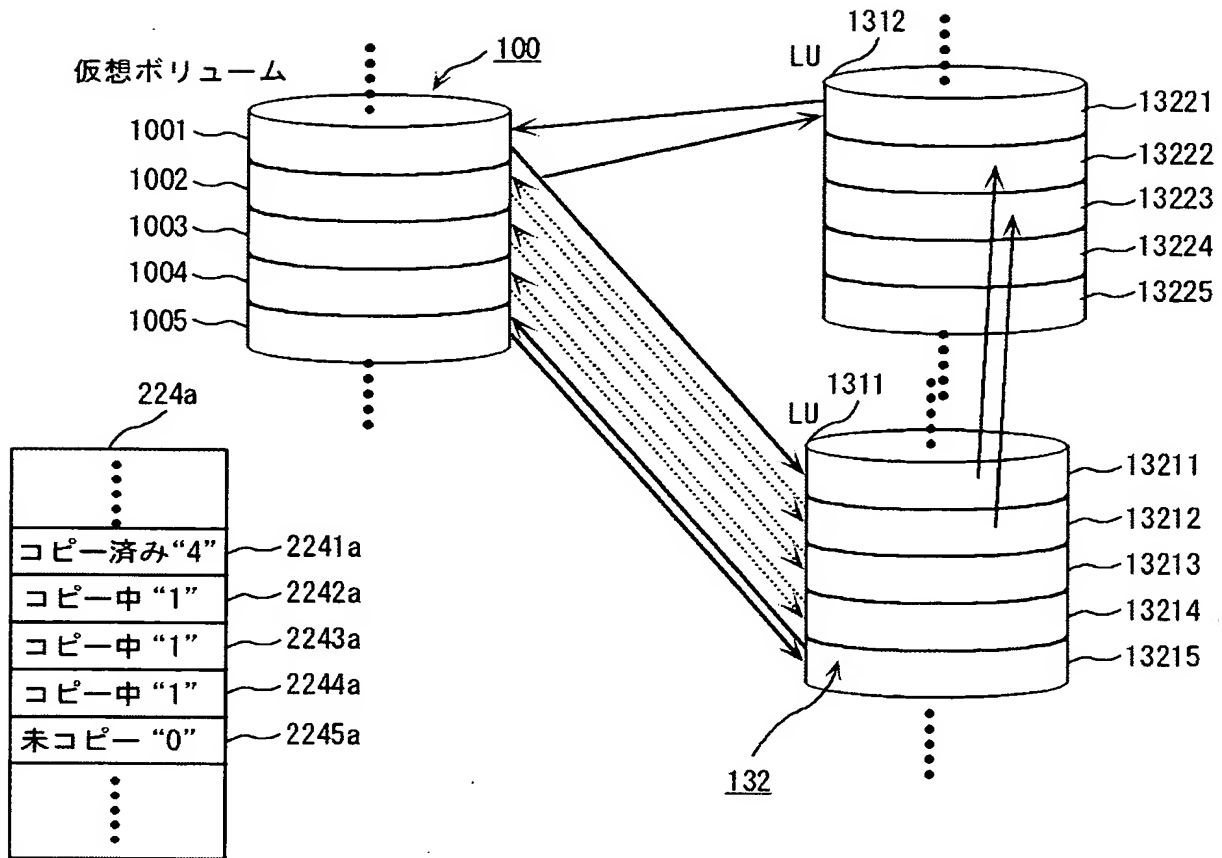
図 25



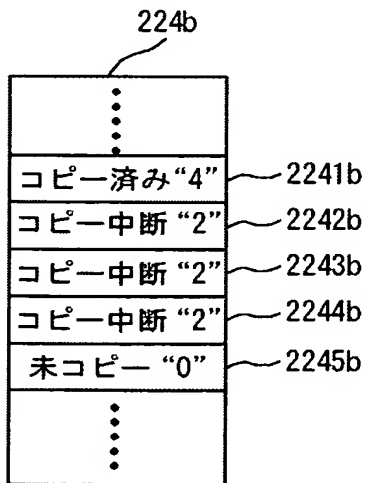
【図 26】

図 26

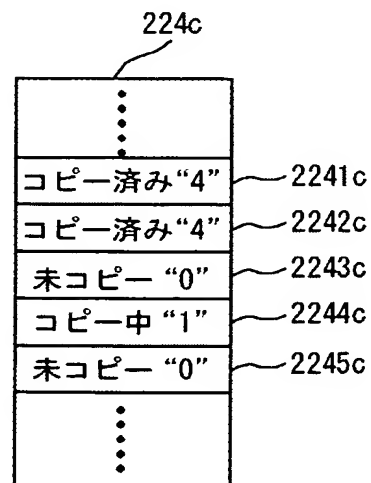
(A)



(B)

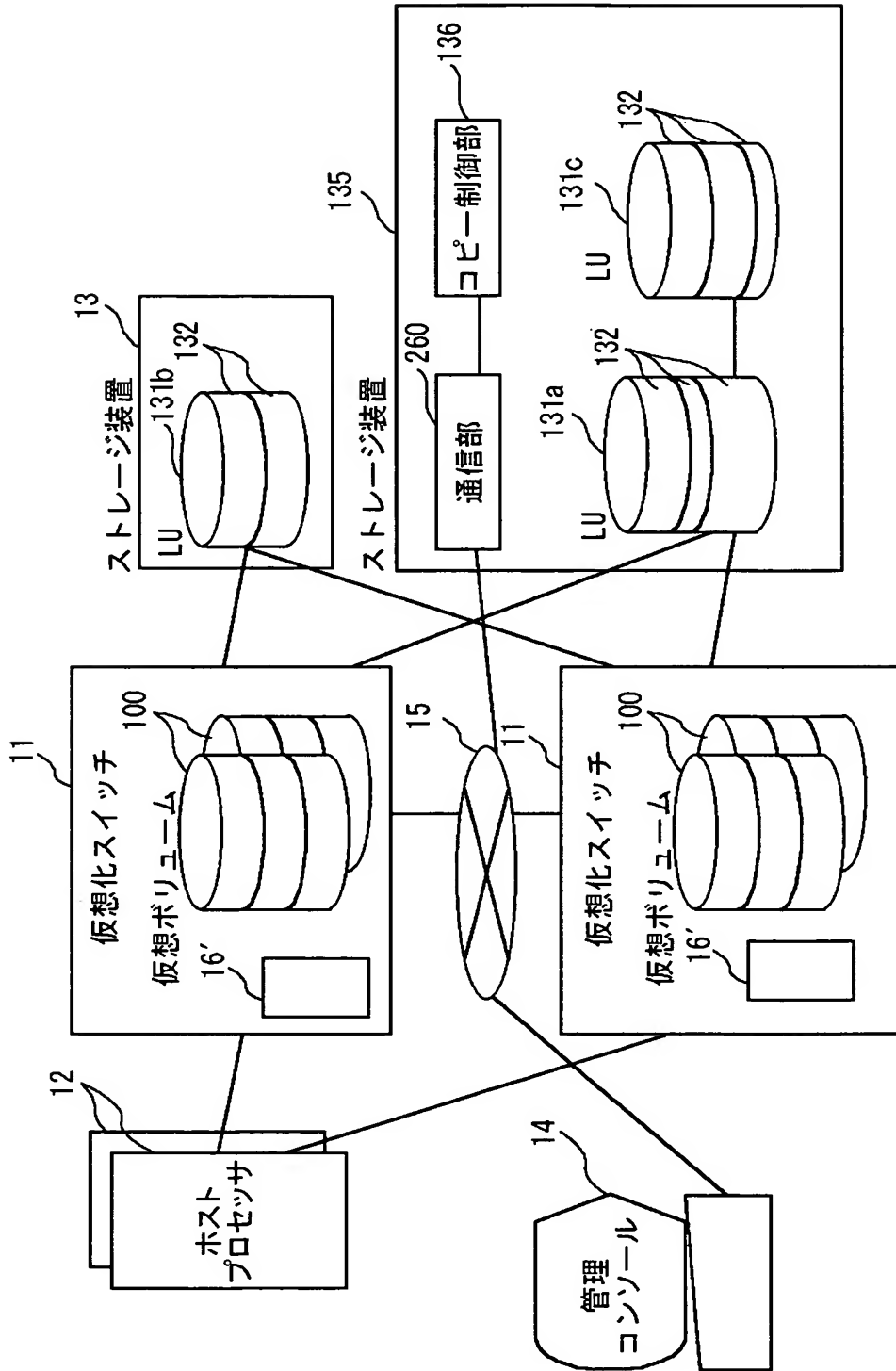


(C)



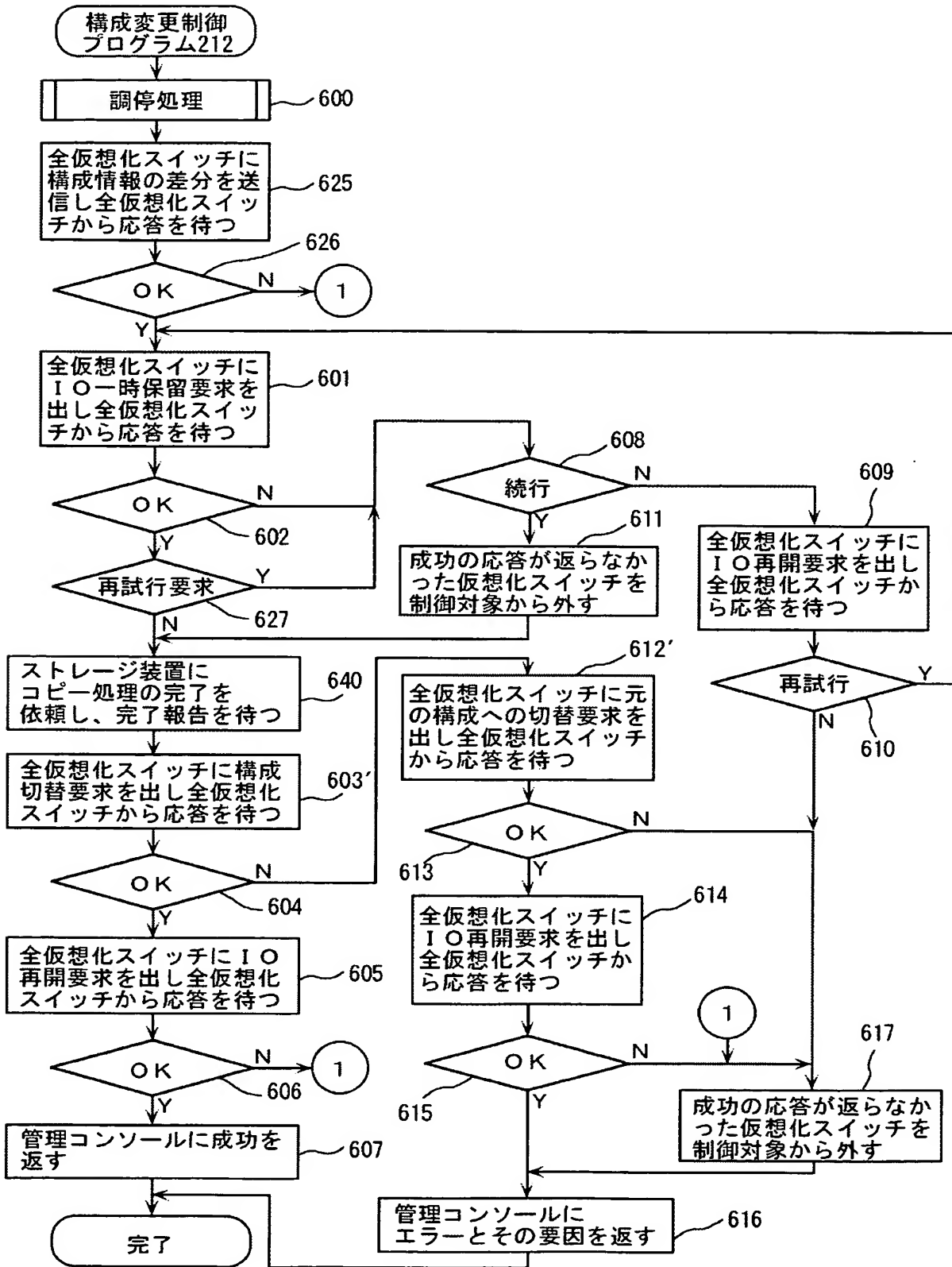
【図 27】

図 27

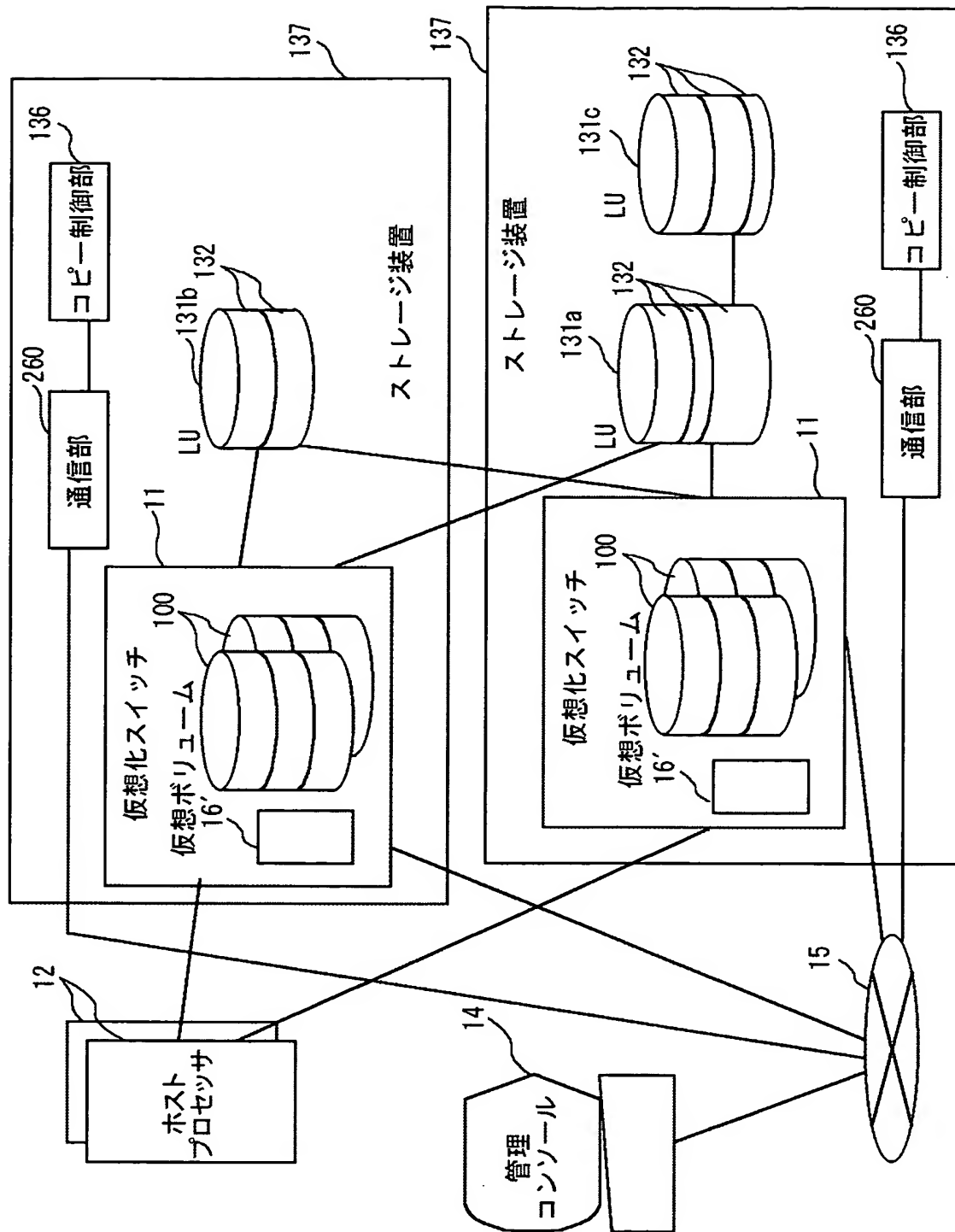
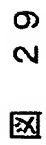


【図 28】

図 28



【圖 29】



【書類名】 要約書**【要約】****【課題】**

冗長構成の仮想化装置を有するストレージシステムにおいて、システムの稼動中にデータ破壊や入出力障害を回避して、ストレージ仮想化装置の構成情報を変更する。

【解決手段】

ストレージ装置が持つある記憶領域を割当てて複数の仮想ボリュームを形成すると共に、ある仮想ボリュームに対するホストプロセッサからの入出力を処理する複数の仮想化装置を有するストレージシステムにおいて、複数の仮想化装置に対し、ホストプロセッサからの受け付けている処理中の入出力を全て完了させ、以降受け付ける入出力の処理を一時保留させるための要求を発行し、この要求に対する処理中の入出力の処理の完了報告を複数の仮想化装置から受け、要求を発した全ての仮想化装置から完了報告を受けた場合、全ての仮想化装置に対してストレージ装置の記憶領域の割り当て変更の指示を送り、全ての仮想化装置からの割り当ての変更の完了報告を受けた場合、一時保留された入出力の状態を解除するための指示を全ての仮想化装置に送ることにより仮想ボリュームに対してストレージ装置の記憶領域の割り当てる。

【選択図】 図 1

特願 2 0 0 3 - 3 4 6 6 0 8

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 1 0 8]

1. 変更年月日

1 9 9 0 年 8 月 3 1 日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台 4 丁目 6 番地

氏 名

株式会社日立製作所